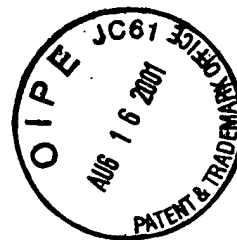


#4



# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.
- (ii) TITLE OF INVENTION: Novel Telomerase.
- (iii) NUMBER OF SEQUENCES: 225
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESS: Townsend and Townsend and Crew LLP
  - (B) STREET: Two Embarcadero Center, 8th Floor
  - (C) CITY: San Francisco
  - (D) STATE: California
  - (E) COUNTRY: United States of America
  - (F) ZIP: 94111
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/854,050
  - (B) FILING DATE: 09-MAY-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/851,843
  - (B) FILING DATE: 06-MAY-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/846,017
  - (B) FILING DATE: 25-APR-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/844,419
  - (B) FILING DATE: 18-APR-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/724,643
  - (B) FILING DATE: 01-OCT-1996
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Apple, Randolph T.
  - (B) REGISTRATION NUMBER: 36,429
  - (C) REFERENCE/DOCKET NUMBER: 015389-002930US
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (415) 576-0200

(B) TELEFAX: (415) 576-0300

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3279 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT AACCTCAGTA	60
TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA ATGGAGGTTG ATGTTGATAA	120
TCAAGCTGAT AATCATGGCA TTCACTCAGC TCTTAAGACT TGTGAAGAAA TTAAAGAAGC	180
TAAAACGTTG TACTCTTGGA TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA	240
TTATAAAGAT TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTIONCACG	300
AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT CAACTGGACT	360
AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT TCATCAAGCG ATGTTTCAGA	420
TAGACAAAAA CTTCAATGAT TTGGATTTC AACTTAAGGGA AATCAATTAG CAAAGACCCA	480
TTTATTAACA GCTCTTTCAA CTCAAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT	540
TAGAGCAATG ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA	600
GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG ATCATTTGAA	660
AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA GCAGACATGA ATGAACCTCG	720
ATGTTGATCA ACCTGCAAAT ACAATGTCAA GAATGAGAAA GATCACTTTC TCAACAACAT	780
CAACGTGCCG AATTGGAATA ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATT	840
TAATAGAAAT AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTC	900
AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA GAATTAGAAA	960
GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG CTTGAGAAAG TCAAAGATTT	1020
TAACCTCAAC TACTATTTAA CAAAATCTTG TCCTCTTCCA GAAAATTGGC GGGAACGGAA	1080
ACAAAAAATC GAAAACCTGA TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA	1140
GCTGTTTAGC TACACAACTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA	1200
CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC AAAAGAAAGT	1260
TAAGAAATAT GTGGAACCTAA ACAAGCATGA ACTCATTAC AAAAACTTAT TGCTTGAGAA	1320
GATCAATACA AGAGAAATAT CATGGATGCA GGTGAGACC TCTGCAAAGC ATTTTTATTA	1380
TTTTGATCAC GAAAACATCT ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT	1440

CGTCGTCTCG CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA	1500
AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA TCGCAGACTT	1560
AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT GAAGAATGGA AAAAGTCGCT	1620
TGGATTTGCA CCTGGAAAAC TCAGACTAAT ACCGAAGAAA ACTACTTTCC GTCCAATTAT	1680
GACTTTCAAT AAGAAGATTG TAAATTCAGA CCGGAAGACT ACAAATTAA CTACAAATAC	1740
GAAGTTATTG AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC	1800
TTTTGGATTG GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG AGTTTGTTTTG	1860
CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTTGCA ACTATGGATA TCGAAAAGTG	1920
ATATGATAGT GTAAACAGAG AAAAATATC AACATTCTTA AAACTACTA AATTACTTTC	1980
TTCAGATTTT TGGATTATGA CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA	2040
TTCGAAAAAC TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT	2100
TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG AACAAAATGA	2160
CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA AGAAATTATT TTAAGAAAGA	2220
TAAGTTACTT CAACCACTCA TTAATATTTT CCAATATAAT TACATTAAGT TTAATGGGAA	2280
GTTTTATAAA CAAACAAAAG GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTCATC	2340
ATTTTATTAT GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA	2400
CCCTGAAAAT CCAAATGTGA ATCTTCTAAT GAGACTTACA GATGACTATC TTTTGATTAC	2460
AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT ATAAACGTAA GTCGTGAAAA	2520
TGGATTTAAA TTCAATATGA AGAAACTACA GACTAGTTTT CCATTAAGTC CAAGCAAATT	2580
TGCAAAATAC GGAATGGATA GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG	2640
GATTGGCATC TCAATTGATA TGAAAACCTT TGCTTTAATG CCAAATATTA ACTTGAGAAT	2700
AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT CAATGTGGCT	2760
CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC CATTATTTTA GAAAGACGAT	2820
TACAACCGAA GACTTTGCGA ATAAAACCTT CAACAAGTTA TTTATATCAG GCGGTTACAA	2880
ATACATGCAA TGAGCCAAAG AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG	2940
TATGATCGAC TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA	3000
CCTTGTTGTC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG ACTTTTTCTT	3060
TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA AAGTACATTT TCAACAGAGT	3120
TTGCATGATC CTCAAGGCAA AAGAAGCAAA GCTAAAAAGT GACCAATGTC AATCTCTAAT	3180
TCAATATGAT GCATAGTCGA CTATTCTAAC TTATTTTGGG AAGTTAATTT TCAATTTTTG	3240
TCTTATATAC TGGGGTTTTG GGGTTTTGGG GTTTTGGGG	3279

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1031 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	Val	Asp	Val	Asp	Asn	Gln	Ala	Asp	Asn	His	Gly	Ile	His	Ser	
1				5					10					15		
Ala	Leu	Lys	Thr	Cys	Glu	Glu	Ile	Lys	Glu	Ala	Lys	Thr	Leu	Tyr	Ser	
			20					25					30			
Trp	Ile	Gln	Lys	Val	Ile	Arg	Cys	Arg	Asn	Gln	Ser	Gln	Ser	His	Tyr	
			35				40					45				
Lys	Asp	Leu	Glu	Asp	Ile	Lys	Ile	Phe	Ala	Gln	Thr	Asn	Ile	Val	Ala	
	50					55					60					
Thr	Pro	Arg	Asp	Tyr	Asn	Glu	Glu	Asp	Phe	Lys	Val	Ile	Ala	Arg	Lys	
65					70					75					80	
Glu	Val	Phe	Ser	Thr	Gly	Leu	Met	Ile	Glu	Leu	Ile	Asp	Lys	Cys	Leu	
				85					90					95		
Val	Glu	Leu	Leu	Ser	Ser	Ser	Asp	Val	Ser	Asp	Arg	Gln	Lys	Leu	Gln	
			100					105					110			
Cys	Phe	Gly	Phe	Gln	Leu	Lys	Gly	Asn	Gln	Leu	Ala	Lys	Thr	His	Leu	
		115					120					125				
Leu	Thr	Ala	Leu	Ser	Thr	Gln	Lys	Gln	Tyr	Phe	Phe	Gln	Asp	Glu	Trp	
	130					135						140				
Asn	Gln	Val	Arg	Ala	Met	Ile	Gly	Asn	Glu	Leu	Phe	Arg	His	Leu	Tyr	
145					150					155					160	
Thr	Lys	Tyr	Leu	Ile	Phe	Gln	Arg	Thr	Ser	Glu	Gly	Thr	Leu	Val	Gln	
			165						170					175		
Phe	Cys	Gly	Asn	Asn	Val	Phe	Asp	His	Leu	Lys	Val	Asn	Asp	Lys	Phe	
			180					185					190			
Asp	Lys	Lys	Gln	Lys	Gly	Gly	Ala	Ala	Asp	Met	Asn	Glu	Pro	Arg	Cys	
		195					200					205				
Cys	Ser	Thr	Cys	Lys	Tyr	Asn	Val	Lys	Asn	Glu	Lys	Asp	His	Phe	Leu	
		210				215					220					
Asn	Asn	Ile	Asn	Val	Pro	Asn	Trp	Asn	Asn	Met	Lys	Ser	Arg	Thr	Arg	
225					230					235					240	
Ile	Phe	Tyr	Cys	Thr	His	Phe	Asn	Arg	Asn	Asn	Gln	Phe	Phe	Lys	Lys	
			245					250						255		
His	Glu	Phe	Val	Ser	Asn	Lys	Asn	Asn	Ile	Ser	Ala	Met	Asp	Arg	Ala	
			260				265						270			

Gln Thr Ile Phe Thr Asn Ile Phe Arg Phe Asn Arg Ile Arg Lys Lys  
 275 280 285  
 Leu Lys Asp Lys Val Ile Glu Lys Ile Ala Tyr Met Leu Glu Lys Val  
 290 295 300  
 Lys Asp Phe Asn Phe Asn Tyr Tyr Leu Thr Lys Ser Cys Pro Leu Pro  
 305 310 315 320  
 Glu Asn Trp Arg Glu Arg Lys Gln Lys Ile Glu Asn Leu Ile Asn Lys  
 325 330 335  
 Thr Arg Glu Glu Lys Ser Lys Tyr Tyr Glu Glu Leu Phe Ser Tyr Thr  
 340 345 350  
 Thr Asp Asn Lys Cys Val Thr Gln Phe Ile Asn Glu Phe Phe Tyr Asn  
 355 360 365  
 Ile Leu Pro Lys Asp Phe Leu Thr Gly Arg Asn Arg Lys Asn Phe Gln  
 370 375 380  
 Lys Lys Val Lys Lys Tyr Val Glu Leu Asn Lys His Glu Leu Ile His  
 385 390 395 400  
 Lys Asn Leu Leu Leu Glu Lys Ile Asn Thr Arg Glu Ile Ser Trp Met  
 405 410 415  
 Gln Val Glu Thr Ser Ala Lys His Phe Tyr Tyr Phe Asp His Glu Asn  
 420 425 430  
 Ile Tyr Val Leu Trp Lys Leu Leu Arg Trp Ile Phe Glu Asp Leu Val  
 435 440 445  
 Val Ser Leu Ile Arg Cys Phe Phe Tyr Val Thr Glu Gln Gln Lys Ser  
 450 455 460  
 Tyr Ser Lys Thr Tyr Tyr Tyr Arg Lys Asn Ile Trp Asp Val Ile Met  
 465 470 475 480  
 Lys Met Ser Ile Ala Asp Leu Lys Lys Glu Thr Leu Ala Glu Val Gln  
 485 490 495  
 Glu Lys Glu Val Glu Glu Trp Lys Lys Ser Leu Gly Phe Ala Pro Gly  
 500 505 510  
 Lys Leu Arg Leu Ile Pro Lys Lys Thr Thr Phe Arg Pro Ile Met Thr  
 515 520 525  
 Phe Asn Lys Lys Ile Val Asn Ser Asp Arg Lys Thr Thr Lys Leu Thr  
 530 535 540  
 Thr Asn Thr Lys Leu Leu Asn Ser His Leu Met Leu Lys Thr Leu Lys  
 545 550 555 560  
 Asn Arg Met Phe Lys Asp Pro Phe Gly Phe Ala Val Phe Asn Tyr Asp  
 565 570 575  
 Asp Val Met Lys Lys Tyr Glu Glu Phe Val Cys Lys Trp Lys Gln Val  
 580 585 590  
 Gly Gln Pro Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr  
 595 600 605

Asp Ser Val Asn Arg Glu Lys Leu Ser Thr Phe Leu Lys Thr Thr Lys  
 610 615 620  
 Leu Leu Ser Ser Asp Phe Trp Ile Met Thr Ala Gln Ile Leu Lys Arg  
 625 630 635 640  
 Lys Asn Asn Ile Val Ile Asp Ser Lys Asn Phe Arg Lys Lys Glu Met  
 645 650 655  
 Lys Asp Tyr Phe Arg Gln Lys Phe Gln Lys Ile Ala Leu Glu Gly Gly  
 660 665 670  
 Gln Tyr Pro Thr Leu Phe Ser Val Leu Glu Asn Glu Gln Asn Asp Leu  
 675 680 685  
 Asn Ala Lys Lys Thr Leu Ile Val Glu Ala Lys Gln Arg Asn Tyr Phe  
 690 695 700  
 Lys Lys Asp Asn Leu Leu Gln Pro Val Ile Asn Ile Cys Gln Tyr Asn  
 705 710 715 720  
 Tyr Ile Asn Phe Asn Gly Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro  
 725 730 735  
 Gln Gly Leu Cys Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr  
 740 745 750  
 Leu Glu Glu Ser Ser Leu Gly Phe Leu Arg Asp Glu Ser Met Asn Pro  
 755 760 765  
 Glu Asn Pro Asn Val Asn Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu  
 770 775 780  
 Leu Ile Thr Thr Gln Glu Asn Asn Ala Val Leu Phe Ile Glu Lys Leu  
 785 790 795 800  
 Ile Asn Val Ser Arg Glu Asn Gly Phe Lys Phe Asn Met Lys Lys Leu  
 805 810 815  
 Gln Thr Ser Phe Pro Leu Ser Pro Ser Lys Phe Ala Lys Tyr Gly Met  
 820 825 830  
 Asp Ser Val Glu Glu Gln Asn Ile Val Gln Asp Tyr Cys Asp Trp Ile  
 835 840 845  
 Gly Ile Ser Ile Asp Met Lys Thr Leu Ala Leu Met Pro Asn Ile Asn  
 850 855 860  
 Leu Arg Ile Glu Gly Ile Leu Cys Thr Leu Asn Leu Asn Met Gln Thr  
 865 870 875 880  
 Lys Lys Ala Ser Met Trp Leu Lys Lys Lys Leu Lys Ser Phe Leu Met  
 885 890 895  
 Asn Asn Ile Thr His Tyr Phe Arg Lys Thr Ile Thr Thr Glu Asp Phe  
 900 905 910  
 Ala Asn Lys Thr Leu Asn Lys Leu Phe Ile Ser Gly Gly Tyr Lys Tyr  
 915 920 925  
 Met Gln Cys Ala Lys Glu Tyr Lys Asp His Phe Lys Lys Asn Leu Ala  
 930 935 940

Met Ser Ser Met Ile Asp Leu Glu Val Ser Lys Ile Ile Tyr Ser Val  
945 950 955 960

Thr Arg Ala Phe Phe Lys Tyr Leu Val Cys Asn Ile Lys Asp Thr Ile  
965 970 975

Phe Gly Glu Glu His Tyr Pro Asp Phe Phe Leu Ser Thr Leu Lys His  
980 985 990

Phe Ile Glu Ile Phe Ser Thr Lys Lys Tyr Ile Phe Asn Arg Val Cys  
995 1000 1005

Met Ile Leu Lys Ala Lys Glu Ala Lys Leu Lys Ser Asp Gln Cys Gln  
1010 1015 1020

Ser Leu Ile Gln Tyr Asp Ala  
1025 1030

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1762 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCCCAAACC CCAAACCCC AAAACCCCTA TAAAAAAGA AAAAATTGAG GTAGTTTAGA	60
AATAAAATAT TATTCCCGCA CAAATGGAGA TGGATATTGA TTTGGATGAT ATAGAAAATT	120
TACTTCCTAA TACATTCAAC AAGTATAGCA GCTCTTGTAG TGACAAGAAA GGATGCAAAA	180
CATTGAAATC TGGCTCGAAA TCGCCTTCAT TGACTATTCC AAAGTTGCAA AAACAATTAG	240
AGTTCTACTT CTCGGATGCA AATCTTTATA ACGATTCTTT CTTGAGAAAA TTAGTTTTAA	300
AAAGCGGAGA GCAAAGAGTA GAAATTGAAA CATTACTAAT GTTTAAATAA AATCAGGTAA	360
TGAGGATTAT TCTATTTTTT AGATCACTTC TTAAGGAGCA TTATGGAGAA AATTACTTAA	420
TACTAAAAGG TAAACAGTTT GGATTATTTT CCTAGCCAAC AATGATGAGT ATATTAAATT	480
CATATGAGAA TGAGTCAAAG GATCTCGATA CATCAGACTT ACCAAAGACA AACTCGCTAT	540
AAAACGCAAG AAAAAGTTTG ATAATCGAAC AGCAGAAGAA CTTATTGCAT TTACTATTCTG	600
TATGGGTTTT ATTACAATTG TTTTAGGTAT CGACGGTGAA CTCCCGAGTC TTGAGACAAT	660
TGAAAAAGCT GTTTACAAC TGAAGGAATCG CAGTTCTGAA AGTTCTGATG TGTATGCCAT	720
TATTTTGTGA ATTAATCTCA AATATCTTAT CTCAATTTAA TGGATAGCTA TAGAAACAAA	780
CCAAATAAAC CATGCAAGTT TAATGGAATA TACGTTAAAT CCTTTGGGAC AAATGCACAC	840
TGAATTTATA TTGGATTCTT AAAGCATAGA TACACAGAAT GCTTTAGAGA CTGATTTAGC	900
TTACAACAGA TTACCTGTTT TGATTACTCT TGCTCATCTC TTATATCTTT AAAAGAAGCA	960

GGCGAAATGA AAAGAAGACT AAAGAAAGAG ATTTCAAAAT TTGTTGATTC TTCTGTAACC 1020  
GGAATTAACA ACAAGAATAT TAGCAACGAA AAAGAAGAAG AGCTATCACA ATCCTGATTC 1080  
TTAAAGATTT CAAAAATTCC AGGTAAGAGA GATACATTCA TTAATAATTCA TATATTATAG 1140  
TTTTTCATTT CACAGCTGTT ATTTTCTTTT ATCTTAACAA TATTTTTTTGA TTAGCTGGAA 1200  
GTAAAAAGTA TCAAATAAGA GAAGCGCTAG ACTGAGGTAA CTTAGCTTAT TCACATTCAT 1260  
AGATCGACCT TCATATATCC AATACGATGA TAAGGAAACA GCAGTCATCC GTTTTAAAAA 1320  
TAGTGCTATG AGGACTAAAT TTTTAGAGTC AAGAAATGGA GCCGAAATCT TAATCAAAAA 1380  
GAATTGCGTC GATATTGCAA AAGAATCGAA CTCTAAATCT TTCGTTAATA AGTATTACCA 1440  
ATCTTGATTG ATTGAAGAGA TTGACGAGGC AACTGCACAG AAGATCATT AAGAAATAAA 1500  
GTAACCTTTTA TTAATTAGAG AATAAACTAA ATTACTAATA TAGAGATCAG CGATCTTCAA 1560  
TTGACGAAAT AAAAGCTGAA CTAAAGTTAG ACAATAAAAA ATACAAACCT TGGTCAAAAT 1620  
ATTGAGGAAG GAAAAGAAGA CCAGTTAGCA AAAGAAAAAA TAAGGCAATA AATAAAATGA 1680  
GTACAGAAAGT GAAGAAATAA AAGATTTATT TTTTCAATA ATTTATTGAA AAGAGGGGTT 1740  
TTGGGGTTTT GGGGTTTTGG GG 1762

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 552 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Pro	Gln	Asn	Pro	Lys	Thr	Pro	Lys	Pro	Leu	Lys	Lys	Lys	Lys	Leu	Arg
1				5					10					15	
Phe	Arg	Asn	Lys	Ile	Leu	Phe	Pro	His	Lys	Trp	Arg	Trp	Ile	Leu	Ile
			20					25					30		
Trp	Met	Ile	Lys	Ile	Tyr	Phe	Leu	Ile	His	Ser	Thr	Ser	Ile	Ala	Ala
			35				40					45			
Leu	Val	Val	Thr	Arg	Lys	Asp	Ala	Lys	His	Cys	Asn	Leu	Ala	Arg	Asn
			50			55					60				
Arg	Leu	His	Cys	Leu	Phe	Gln	Ser	Cys	Lys	Asn	Asn	Ser	Ser	Thr	Ser
65				70						75				80	
Arg	Met	Gln	Ile	Phe	Ile	Thr	Ile	Leu	Ser	Cys	Glu	Asn	Phe	Lys	Ala
			85					90						95	
Glu	Ser	Lys	Glu	Lys	Leu	Lys	His	Tyr	Cys	Leu	Asn	Lys	Ile	Arg	Cys
			100					105					110		
Gly	Leu	Phe	Tyr	Phe	Leu	Asp	His	Phe	Leu	Arg	Ser	Ile	Met	Glu	Lys



115					120					125					
Ile	Thr	Tyr	Lys	Val	Asn	Ser	Leu	Asp	Tyr	Phe	Pro	Ser	Gln	Gln	Cys
	130					135					140				
Cys	Val	Tyr	Ile	His	Met	Arg	Met	Ser	Gln	Arg	Ile	Ser	Ile	His	Gln
145					150					155					160
Thr	Tyr	Gln	Arg	Gln	Thr	Arg	Tyr	Lys	Thr	Gln	Glu	Lys	Val	Cys	Ser
				165					170					175	
Asn	Ser	Arg	Arg	Thr	Tyr	Cys	Ile	Tyr	Tyr	Ser	Tyr	Gly	Phe	Tyr	Tyr
			180					185					190		
Asn	Cys	Phe	Arg	Tyr	Arg	Arg	Cys	Thr	Pro	Glu	Ser	Cys	Asp	Asn	Cys
		195					200					205			
Lys	Ser	Cys	Leu	Gln	Leu	Lys	Glu	Ser	Gln	Phe	Cys	Lys	Phe	Cys	Val
	210					215					220				
Cys	His	Tyr	Phe	Val	Asn	Ser	Gln	Ile	Ser	Tyr	Leu	Asn	Leu	Met	Asp
225						230					235				240
Ser	Tyr	Arg	Asn	Lys	Pro	Asn	Lys	Pro	Cys	Lys	Phe	Asn	Gly	Ile	Tyr
				245					250					255	
Val	Lys	Ser	Phe	Gly	Thr	Asn	Ala	His	Cys	Ile	Tyr	Ile	Gly	Phe	Leu
			260					265					270		
Lys	His	Arg	Tyr	Thr	Glu	Cys	Phe	Arg	Asp	Cys	Phe	Ser	Leu	Gln	Gln
		275					280					285			
Ile	Thr	Cys	Phe	Asp	Tyr	Ser	Cys	Ser	Ser	Leu	Ile	Ser	Leu	Lys	Glu
	290					295					300				
Ala	Gly	Glu	Met	Lys	Arg	Arg	Leu	Lys	Lys	Glu	Ile	Ser	Lys	Phe	Val
305						310					315				320
Asp	Ser	Ser	Val	Thr	Gly	Ile	Asn	Asn	Lys	Asn	Ile	Ser	Asn	Glu	Lys
				325					330					335	
Glu	Glu	Glu	Leu	Ser	Gln	Ser	Cys	Phe	Leu	Lys	Ile	Ser	Lys	Ile	Pro
			340					345					350		
Gly	Lys	Arg	Asp	Thr	Phe	Ile	Lys	Ile	His	Ile	Leu	Phe	Phe	Ile	Ser
		355					360					365			
Gln	Leu	Leu	Phe	Ser	Phe	Ile	Leu	Thr	Ile	Phe	Phe	Asp	Leu	Glu	Val
	370					375					380				
Lys	Ser	Ile	Lys	Glu	Lys	Arg	Thr	Glu	Val	Thr	Leu	Ile	His	Ile	His
385						390					395				400
Arg	Ser	Thr	Phe	Ile	Tyr	Pro	Ile	Arg	Cys	Gly	Asn	Ser	Ser	His	Pro
				405					410					415	
Phe	Lys	Cys	Tyr	Glu	Asp	Ile	Phe	Arg	Val	Lys	Lys	Trp	Ser	Arg	Asn
			420					425					430		
Leu	Asn	Gln	Lys	Glu	Leu	Arg	Arg	Tyr	Cys	Lys	Arg	Ile	Glu	Leu	Ile
		435					440					445			

Phe Arg Val Leu Pro Ile Leu Ile Asp Cys Arg Asp Arg Gly Asn Cys  
 450 455 460  
 Thr Glu Asp His Arg Asn Lys Val Thr Phe Ile Asn Arg Ile Asn Ile  
 465 470 475 480  
 Thr Asn Ile Glu Ile Ser Asp Leu Gln Leu Thr Lys Lys Leu Asn Ser  
 485 490 495  
 Thr Ile Lys Asn Thr Asn Leu Gly Gln Asn Ile Glu Glu Gly Lys Glu  
 500 505 510  
 Asp Gln Leu Ala Lys Glu Lys Ile Arg Gln Ile Lys Cys Val Gln Lys  
 515 520 525  
 Cys Arg Asn Lys Arg Phe Ile Phe Phe Asn Asn Leu Leu Lys Arg Gly  
 530 535 540  
 Val Leu Gly Phe Trp Gly Phe Gly  
 545 550

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 562 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Pro Lys Thr Pro Lys Pro Gln Asn Pro Tyr Lys Lys Arg Lys Asn Cys  
 1 5 10 15  
 Gly Ser Leu Glu Ile Lys Tyr Tyr Ser Arg Thr Asn Gly Asp Gly Tyr  
 20 25 30  
 Cys Phe Gly Cys Tyr Arg Lys Phe Thr Ser Tyr Ile Gln Val Gln  
 35 40 45  
 Leu Leu Gln Glu Arg Met Gln Asn Ile Glu Ile Trp Leu Glu Ile Ala  
 50 55 60  
 Phe Ile Asp Tyr Ser Lys Val Ala Lys Thr Ile Arg Val Leu Leu Leu  
 65 70 75 80  
 Gly Cys Lys Ser Leu Arg Phe Phe Leu Glu Lys Ile Ser Phe Lys Lys  
 85 90 95  
 Arg Arg Ala Lys Ser Arg Asn Cys Asn Ile Thr Asn Val Ile Lys Ser  
 100 105 110  
 Gly Asn Glu Asp Tyr Ser Ile Phe Ile Thr Ser Gly Ala Leu Trp Arg  
 115 120 125  
 Lys Leu Leu Asn Thr Lys Arg Thr Val Trp Ile Ile Ser Leu Ala Asn  
 130 135 140  
 Asn Asp Glu Tyr Ile Lys Phe Ile Cys Glu Cys Val Lys Gly Ser Arg  
 145 150 155 160

Tyr Ile Arg Leu Thr Lys Asp Lys Leu Ala Ile Lys Arg Lys Lys Lys  
165 170 175  
Phe Asp Asn Arg Thr Ala Glu Glu Leu Ile Ala Phe Thr Ile Arg Met  
180 185 190  
Gly Phe Ile Thr Ile Val Leu Gly Ile Asp Gly Glu Leu Pro Ser Leu  
195 200 205  
Glu Thr Ile Glu Lys Ala Val Tyr Asn Cys Arg Asn Arg Ser Ser Glu  
210 215 220  
Ser Ser Asp Val Tyr Ala Ile Ile Leu Cys Ile Asn Leu Lys Tyr Leu  
225 230 235 240  
Ile Ser Ile Trp Ile Ala Ile Glu Thr Asn Gln Ile Asn His Ala Ser  
245 250 255  
Leu Met Glu Tyr Thr Leu Asn Pro Leu Gly Gln Met His Thr Glu Phe  
260 265 270  
Ile Leu Asp Ser Ser Ile Asp Thr Gln Asn Ala Leu Glu Thr Asp Leu  
275 280 285  
Ala Tyr Asn Arg Leu Pro Val Leu Ile Thr Leu Ala His Leu Leu Tyr  
290 295 300  
Leu Lys Lys Gln Ala Lys Cys Lys Glu Asp Arg Lys Arg Phe Gln Asn  
305 310 315 320  
Leu Leu Ile Leu Leu Pro Glu Leu Thr Thr Arg Ile Leu Ala Thr Lys  
325 330 335  
Lys Lys Lys Ser Tyr His Asn Pro Asp Ser Arg Phe Gln Lys Phe Gln  
340 345 350  
Val Arg Glu Ile His Ser Leu Lys Phe Ile Tyr Tyr Ser Phe Ser Phe  
355 360 365  
His Ser Cys Tyr Phe Leu Leu Ser Gln Tyr Phe Leu Ile Ser Trp Lys  
370 375 380  
Lys Val Ser Asn Lys Arg Ser Ala Arg Leu Arg Leu Ser Leu Phe Thr  
385 390 395 400  
Phe Ile Asp Arg Pro Ser Tyr Ile Gln Tyr Asp Asp Lys Glu Thr Ala  
405 410 415  
Val Ile Arg Phe Lys Asn Ser Ala Met Arg Thr Lys Phe Leu Glu Ser  
420 425 430  
Arg Asn Gly Ala Glu Ile Leu Ile Lys Lys Asn Cys Val Asp Ile Ala  
435 440 445  
Lys Glu Ser Asn Ser Lys Ser Phe Val Asn Lys Tyr Tyr Gln Ser Cys  
450 455 460  
Leu Ile Glu Glu Ile Asp Glu Ala Thr Ala Gln Lys Ile Ile Lys Glu  
465 470 475 480  
Ile Lys Leu Leu Leu Ile Arg Glu Thr Lys Leu Leu Ile Arg Ser Ala  
485 490 495

Ile Phe Asn Cys Arg Asn Lys Ser Cys Thr Lys Val Arg Gln Lys Ile  
500 505 510

Gln Thr Leu Val Lys Ile Leu Arg Lys Glu Lys Lys Thr Ser Gln Lys  
515 520 525

Lys Lys Gly Asn Lys Asn Glu Tyr Arg Ser Glu Glu Ile Lys Asp Leu  
530 535 540

Phe Phe Ser Ile Ile Tyr Cys Lys Glu Gly Phe Trp Gly Phe Gly Val  
545 550 555 560

Leu Gly

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 560 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Pro Lys Pro Gln Asn Pro Lys Thr Pro Ile Lys Lys Glu Lys Ile Glu  
1 5 10 15

Val Val Lys Asn Ile Ile Pro Ala Gln Met Glu Met Asp Ile Asp Leu  
20 25 30

Asp Asp Ile Glu Asn Leu Leu Pro Asn Thr Phe Asn Lys Tyr Ser Ser  
35 40 45

Ser Cys Ser Asp Lys Lys Gly Cys Lys Thr Leu Lys Ser Gly Ser Lys  
50 55 60

Ser Pro Ser Leu Thr Ile Pro Lys Leu Gln Lys Gln Leu Glu Phe Tyr  
65 70 75 80

Phe Ser Asp Ala Asn Leu Tyr Asn Asp Ser Phe Leu Arg Lys Leu Val  
85 90 95

Leu Lys Ser Gly Glu Gln Arg Val Glu Ile Glu Thr Leu Leu Met Phe  
100 105 110

Lys Asn Gln Val Met Arg Ile Ile Leu Phe Phe Arg Ser Leu Leu Lys  
115 120 125

Glu His Tyr Gly Glu Asn Tyr Leu Ile Leu Lys Gly Lys Gln Phe Gly  
130 135 140

Leu Phe Pro Pro Thr Met Met Ser Ile Leu Asn Ser Tyr Glu Asn Glu  
145 150 155 160

Ser Lys Asp Leu Asp Thr Ser Asp Leu Pro Lys Thr Asn Ser Leu Asn  
165 170 175

Ala Arg Lys Ser Leu Ile Ile Glu Gln Lys Asn Leu Leu His Leu  
180 185 190

Leu Phe Val Trp Val Leu Leu Gln Leu Phe Val Ser Thr Val Asn Ser  
 195 200 205  
 Arg Val Leu Arg Gln Leu Lys Lys Leu Phe Thr Thr Glu Gly Ile Ala  
 210 215 220  
 Val Leu Lys Val Leu Met Cys Met Pro Leu Phe Cys Glu Leu Ile Ser  
 225 230 235 240  
 Asn Ile Leu Ser Gln Phe Asn Gly Leu Lys Gln Thr Lys Thr Met Gln  
 245 250 255  
 Val Trp Asn Ile Arg Ile Leu Trp Asp Lys Cys Thr Leu Asn Leu Tyr  
 260 265 270  
 Trp Ile Leu Lys Ala Ile His Arg Met Leu Arg Leu Ile Leu Thr Thr  
 275 280 285  
 Asp Tyr Leu Phe Cys Leu Leu Leu Leu Ile Ser Tyr Ile Phe Lys Arg  
 290 295 300  
 Ser Arg Arg Asn Glu Lys Lys Thr Lys Glu Arg Asp Phe Lys Ile Cys  
 305 310 315 320  
 Cys Phe Phe Cys Asn Arg Asn Gln Gln Glu Tyr Gln Arg Lys Arg Arg  
 325 330 335  
 Arg Ala Ile Thr Ile Leu Ile Leu Lys Asp Phe Lys Asn Ser Arg Glu  
 340 345 350  
 Arg Tyr Ile His Asn Ser Tyr Ile Ile Val Phe His Phe Thr Ala Val  
 355 360 365  
 Ile Phe Phe Tyr Leu Asn Asn Ile Phe Cys Leu Ala Gly Ser Lys Lys  
 370 375 380  
 Tyr Gln Ile Arg Glu Ala Leu Asp Cys Gly Asn Leu Ala Tyr Ser His  
 385 390 395 400  
 Ser Ile Asp Leu His Ile Ser Asn Thr Met Ile Arg Lys Gln Gln Ser  
 405 410 415  
 Ser Val Leu Lys Ile Val Leu Cys Gly Leu Asn Phe Ser Gln Glu Met  
 420 425 430  
 Glu Pro Lys Ser Ser Lys Arg Ile Ala Ser Ile Leu Gln Lys Asn Arg  
 435 440 445  
 Thr Leu Asn Leu Ser Leu Ile Ser Ile Thr Asn Leu Asp Cys Leu Lys  
 450 455 460  
 Arg Leu Thr Arg Gln Leu His Arg Arg Ser Leu Lys Lys Ser Asn Phe  
 465 470 475 480  
 Tyr Leu Glu Asn Lys Leu Asn Tyr Tyr Arg Asp Gln Arg Ser Ser Ile  
 485 490 495  
 Asp Glu Ile Lys Ala Glu Leu Lys Leu Asp Asn Lys Lys Tyr Lys Pro  
 500 505 510  
 Trp Ser Lys Tyr Cys Gly Arg Lys Arg Arg Pro Val Ser Lys Arg Lys  
 515 520 525

Asn Lys Ala Ile Asn Lys Met Ser Thr Glu Val Lys Lys Lys Ile Tyr  
530 535 540

Phe Phe Gln Phe Ile Glu Lys Arg Gly Phe Gly Val Leu Gly Phe Trp  
545 550 555 560

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 719 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Glu Ile Glu Asn Asn Gln Ala Gln Gln Pro Lys Ala Glu Lys Leu  
1 5 10 15

Trp Trp Glu Leu Glu Leu Glu Met Gln Glu Asn Gln Asn Asp Ile Gln  
20 25 30

Val Arg Val Lys Ile Asp Asp Pro Lys Gln Tyr Leu Val Asn Val Thr  
35 40 45

Ala Ala Cys Leu Leu Gln Glu Gly Ser Tyr Tyr Gln Asp Lys Asp Glu  
50 55 60

Arg Arg Tyr Ile Ile Thr Lys Ala Leu Leu Glu Val Ala Glu Ser Asp  
65 70 75 80

Pro Glu Phe Ile Cys Gln Leu Ala Val Tyr Ile Arg Asn Glu Leu Tyr  
85 90 95

Ile Arg Thr Thr Thr Asn Tyr Ile Val Ala Phe Cys Val Val His Lys  
100 105 110

Asn Thr Gln Pro Phe Ile Glu Lys Tyr Phe Asn Lys Ala Val Leu Leu  
115 120 125

Pro Asn Asp Leu Leu Glu Val Cys Glu Phe Ala Gln Val Leu Tyr Ile  
130 135 140

Phe Asp Ala Thr Glu Phe Lys Asn Leu Tyr Leu Asp Arg Ile Leu Ser  
145 150 155 160

Gln Asp Ile Arg Lys Glu Leu Thr Phe Arg Lys Cys Leu Gln Arg Cys  
165 170 175

Val Arg Ser Lys Phe Ser Glu Phe Asn Glu Tyr Gln Leu Gly Lys Tyr  
180 185 190

Cys Thr Glu Ser Gln Arg Lys Lys Thr Met Phe Arg Tyr Leu Ser Val  
195 200 205

Thr Asn Lys Gln Lys Trp Asp Gln Thr Lys Lys Lys Arg Lys Glu Asn  
210 215 220

Leu Leu Thr Lys Leu Gln Ala Ile Lys Glu Ser Glu Asp Lys Ser Lys  
225 230 235 240

Arg Glu Thr Gly Asp Ile Met Asn Val Glu Asp Ala Ile Lys Ala Leu  
 245 250 255  
 Lys Pro Ala Val Met Lys Lys Ile Ala Lys Arg Gln Asn Ala Met Lys  
 260 265 270  
 Lys His Met Lys Ala Pro Lys Ile Pro Asn Ser Thr Leu Glu Ser Lys  
 275 280 285  
 Tyr Leu Thr Phe Lys Asp Leu Ile Lys Phe Cys His Ile Ser Glu Pro  
 290 295 300  
 Lys Glu Arg Val Tyr Lys Ile Leu Gly Lys Lys Tyr Pro Lys Thr Glu  
 305 310 315 320  
 Glu Glu Tyr Lys Ala Ala Phe Gly Asp Ser Ala Ser Ala Pro Phe Asn  
 325 330 335  
 Pro Glu Leu Ala Gly Lys Arg Met Lys Ile Glu Ile Ser Lys Thr Trp  
 340 345 350  
 Glu Asn Glu Leu Ser Ala Lys Gly Asn Thr Ala Glu Val Trp Asp Asn  
 355 360 365  
 Leu Ile Ser Ser Asn Gln Leu Pro Tyr Met Ala Met Leu Arg Asn Leu  
 370 375 380  
 Ser Asn Ile Leu Lys Ala Gly Val Ser Asp Thr Thr His Ser Ile Val  
 385 390 395 400  
 Ile Asn Lys Ile Cys Glu Pro Lys Ala Val Glu Asn Ser Lys Met Phe  
 405 410 415  
 Pro Leu Gln Phe Phe Ser Ala Ile Glu Ala Val Asn Glu Ala Val Thr  
 420 425 430  
 Lys Gly Phe Lys Ala Lys Lys Arg Glu Asn Met Asn Leu Lys Gly Gln  
 435 440 445  
 Ile Glu Ala Val Lys Glu Val Val Glu Lys Thr Asp Glu Glu Lys Lys  
 450 455 460  
 Asp Met Glu Leu Glu Gln Thr Glu Glu Gly Glu Phe Val Lys Val Asn  
 465 470 475 480  
 Glu Gly Ile Gly Lys Gln Tyr Ile Asn Ser Ile Glu Leu Ala Ile Lys  
 485 490 495  
 Ile Ala Val Asn Lys Asn Leu Asp Glu Ile Lys Gly His Thr Ala Ile  
 500 505 510  
 Phe Ser Asp Val Ser Gly Ser Met Ser Thr Ser Met Ser Gly Gly Ala  
 515 520 525  
 Lys Lys Tyr Gly Ser Val Arg Thr Cys Leu Glu Cys Ala Leu Val Leu  
 530 535 540  
 Gly Leu Met Val Lys Gln Arg Cys Glu Lys Ser Ser Phe Tyr Ile Phe  
 545 550 555 560  
 Ser Ser Pro Ser Ser Gln Cys Asn Lys Cys Tyr Leu Glu Val Asp Leu  
 565 570 575

Pro Gly Asp Glu Leu Arg Pro Ser Met Gln Lys Leu Leu Gln Glu Lys  
 580 585 590  
 Gly Lys Leu Gly Gly Gly Thr Asp Phe Pro Tyr Glu Cys Ile Asp Glu  
 595 600 605  
 Trp Thr Lys Asn Lys Thr His Val Asp Asn Ile Val Ile Leu Ser Asp  
 610 615 620  
 Met Met Ile Ala Glu Gly Tyr Ser Asp Ile Asn Val Arg Gly Ser Ser  
 625 630 635 640  
 Ile Val Asn Ser Ile Lys Lys Tyr Lys Asp Glu Val Asn Pro Asn Ile  
 645 650 655  
 Lys Ile Phe Ala Val Asp Leu Glu Gly Tyr Gly Lys Cys Leu Asn Leu  
 660 665 670  
 Gly Asp Glu Phe Asn Glu Asn Asn Tyr Ile Lys Ile Phe Gly Met Ser  
 675 680 685  
 Asp Ser Ile Leu Lys Phe Ile Ser Ala Lys Gln Gly Gly Ala Asn Met  
 690 695 700  
 Val Glu Val Ile Lys Asn Phe Ala Leu Gln Lys Ile Gly Gln Lys  
 705 710 715

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 872 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Ser Arg Arg Asn Gln Lys Lys Pro Gln Ala Pro Ile Gly Asn Glu  
 1 5 10 15  
 Thr Asn Leu Asp Phe Val Leu Gln Asn Leu Glu Val Tyr Lys Ser Gln  
 20 25 30  
 Ile Glu His Tyr Lys Thr Gln Gln Gln Gln Ile Lys Glu Glu Asp Leu  
 35 40 45  
 Lys Leu Leu Lys Phe Lys Asn Gln Asp Gln Asp Gly Asn Ser Gly Asn  
 50 55 60  
 Asp Asp Asp Asp Glu Glu Asn Asn Ser Asn Lys Gln Gln Glu Leu Leu  
 65 70 75 80  
 Arg Arg Val Asn Gln Ile Lys Gln Gln Val Gln Leu Ile Lys Lys Val  
 85 90 95  
 Gly Ser Lys Val Glu Lys Asp Leu Asn Leu Asn Glu Asp Glu Asn Lys  
 100 105 110  
 Lys Asn Gly Leu Ser Glu Gln Gln Val Lys Glu Glu Gln Leu Arg Thr  
 115 120 125



Ile	Thr	Glu	Glu	Gln	Val	Lys	Tyr	Gln	Asn	Leu	Val	Phe	Asn	Met	Asp		
130						135					140						
Tyr	Gln	Leu	Asp	Leu	Asn	Glu	Ser	Gly	Gly	His	Arg	Arg	His	Arg	Arg		
145					150					155						160	
Glu	Thr	Asp	Tyr	Asp	Thr	Glu	Lys	Trp	Phe	Glu	Ile	Ser	His	Asp	Gln		
				165					170					175			
Lys	Asn	Tyr	Val	Ser	Ile	Tyr	Ala	Asn	Gln	Lys	Thr	Ser	Tyr	Cys	Trp		
			180					185					190				
Trp	Leu	Lys	Asp	Tyr	Phe	Asn	Lys	Asn	Asn	Tyr	Asp	His	Leu	Asn	Val		
		195					200					205					
Ser	Ile	Asn	Arg	Leu	Glu	Thr	Glu	Ala	Glu	Phe	Tyr	Ala	Phe	Asp	Asp		
	210					215					220						
Phe	Ser	Gln	Thr	Ile	Lys	Leu	Thr	Asn	Asn	Ser	Tyr	Gln	Thr	Val	Asn		
225					230					235					240		
Ile	Asp	Val	Asn	Phe	Asp	Asn	Asn	Leu	Cys	Ile	Leu	Ala	Leu	Leu	Arg		
				245					250					255			
Phe	Leu	Leu	Ser	Leu	Glu	Arg	Phe	Asn	Ile	Leu	Asn	Ile	Arg	Ser	Ser		
			260					265					270				
Tyr	Thr	Arg	Asn	Gln	Tyr	Asn	Phe	Glu	Lys	Ile	Gly	Glu	Leu	Leu	Glu		
		275					280					285					
Thr	Ile	Phe	Ala	Val	Val	Phe	Ser	His	Arg	His	Leu	Gln	Gly	Ile	His		
	290					295					300						
Leu	Gln	Val	Pro	Cys	Glu	Ala	Phe	Gln	Tyr	Leu	Val	Asn	Ser	Ser	Ser		
305					310					315					320		
Gln	Ile	Ser	Val	Lys	Asp	Ser	Gln	Leu	Gln	Val	Tyr	Ser	Phe	Ser	Thr		
				325					330					335			
Asp	Leu	Lys	Leu	Val	Asp	Thr	Asn	Lys	Val	Gln	Asp	Tyr	Phe	Lys	Phe		
			340					345					350				
Leu	Gln	Glu	Phe	Pro	Arg	Leu	Thr	His	Val	Ser	Gln	Gln	Ala	Ile	Pro		
		355					360					365					
Val	Ser	Ala	Thr	Asn	Ala	Val	Glu	Asn	Leu	Asn	Val	Leu	Leu	Lys	Lys		
	370					375					380						
Val	Lys	His	Ala	Asn	Leu	Asn	Leu	Val	Ser	Ile	Pro	Thr	Gln	Phe	Asn		
385					390					395					400		
Phe	Asp	Phe	Tyr	Phe	Val	Asn	Leu	Gln	His	Leu	Lys	Leu	Glu	Phe	Gly		
				405				410						415			
Leu	Glu	Pro	Asn	Ile	Leu	Thr	Lys	Gln	Lys	Leu	Glu	Asn	Leu	Leu	Leu		
			420					425					430				
Ser	Ile	Lys	Gln	Ser	Lys	Asn	Leu	Lys	Phe	Leu	Arg	Leu	Asn	Phe	Tyr		
		435					440					445					
Thr	Tyr	Val	Ala	Gln	Glu	Thr	Ser	Arg	Lys	Gln	Ile	Leu	Lys	Gln	Ala		
	450					455					460						

Thr Thr Ile Lys Asn Leu Lys Asn Asn Lys Asn Gln Glu Glu Thr Pro  
 465 470 475 480  
 Glu Thr Lys Asp Glu Thr Pro Ser Glu Ser Thr Ser Gly Met Lys Phe  
 485 490 495  
 Phe Asp His Leu Ser Glu Leu Thr Glu Leu Glu Asp Phe Ser Val Asn  
 500 505 510  
 Leu Gln Ala Thr Gln Glu Ile Tyr Asp Ser Leu His Lys Leu Leu Ile  
 515 520 525  
 Arg Ser Thr Asn Leu Lys Lys Phe Lys Leu Ser Tyr Lys Tyr Glu Met  
 530 535 540  
 Glu Lys Ser Lys Met Asp Thr Phe Ile Asp Leu Lys Asn Ile Tyr Glu  
 545 550 555 560  
 Thr Leu Asn Asn Leu Lys Arg Cys Ser Val Asn Ile Ser Asn Pro His  
 565 570 575  
 Gly Asn Ile Ser Tyr Glu Leu Thr Asn Lys Asp Ser Thr Phe Tyr Lys  
 580 585 590  
 Phe Lys Leu Thr Leu Asn Gln Glu Leu Gln His Ala Lys Tyr Thr Phe  
 595 600 605  
 Lys Gln Asn Glu Phe Gln Phe Asn Asn Val Lys Ser Ala Lys Ile Glu  
 610 615 620  
 Ser Ser Ser Leu Glu Ser Leu Glu Asp Ile Asp Ser Leu Cys Lys Ser  
 625 630 635 640  
 Ile Ala Ser Cys Lys Asn Leu Gln Asn Val Asn Ile Ile Ala Ser Leu  
 645 650 655  
 Leu Tyr Pro Asn Asn Ile Gln Lys Asn Pro Phe Asn Lys Pro Asn Leu  
 660 665 670  
 Leu Phe Phe Lys Gln Phe Glu Gln Leu Lys Asn Leu Glu Asn Val Ser  
 675 680 685  
 Ile Asn Cys Ile Leu Asp Gln His Ile Leu Asn Ser Ile Ser Glu Phe  
 690 695 700  
 Leu Glu Lys Asn Lys Lys Ile Lys Ala Phe Ile Leu Lys Arg Tyr Tyr  
 705 710 715 720  
 Leu Leu Gln Tyr Tyr Leu Asp Tyr Thr Lys Leu Phe Lys Thr Leu Gln  
 725 730 735  
 Gln Leu Pro Glu Leu Asn Gln Val Tyr Ile Asn Gln Gln Leu Glu Glu  
 740 745 750  
 Leu Thr Val Ser Glu Val His Lys Gln Val Trp Glu Asn His Lys Gln  
 755 760 765  
 Lys Ala Phe Tyr Glu Pro Leu Cys Glu Phe Ile Lys Glu Ser Ser Gln  
 770 775 780  
 Thr Leu Gln Leu Ile Asp Phe Asp Gln Asn Thr Val Ser Asp Asp Ser  
 785 790 795 800





Ser Met Ser Gly Gly Ala Lys Lys Tyr Gly Ser Val Arg Thr Cys Leu  
35 40 45

Glu Cys Ala Leu Val Leu Gly Leu Met Val Lys Gln Arg Cys Glu Lys  
50 55 60

Ser Ser Phe Tyr Ile Phe Ser Ser Pro Ser Ser Gln Cys Lys Cys Tyr  
65 70 75 80

Leu Glu Val Asp Leu Pro Gly Asp Glu Leu Arg Pro Ser Met Gln Lys  
85 90 95

Leu Leu

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 69 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Gln Pro Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr  
1 5 10 15

Asp Ser Val Asn Arg Glu Lys Leu Ser Thr Phe Leu Lys Thr Thr Lys  
20 25 30

Leu Leu Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro Gln Gly Leu Cys  
35 40 45

Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr Leu Glu Glu Ser  
50 55 60

Ser Leu Gly Phe Leu  
65

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 69 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Lys Asn Arg Asn Leu His Cys Thr Tyr Ile Asp Tyr Lys Lys Ala Phe  
1 5 10 15

Asp Ser Ile Pro His Ser Trp Leu Ile Gln Val Leu Glu Ile Tyr Lys  
20 25 30

Ile Asn Arg Gln Ile Ala Ile Lys Gly Ile Tyr Gln Gly Asp Ser  
35 40 45

Leu Ser Pro Leu Trp Phe Cys Leu Ala Leu Asn Pro Leu Ser His Gln  
 50 55 60

Leu His Asn Asp Arg  
 65

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 69 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Phe Gly Gly Ser Asn Trp Phe Arg Glu Val Asp Leu Lys Lys Cys Phe  
 1 5 10 15  
 Asp Thr Ile Ser His Asp Leu Ile Ile Lys Glu Leu Lys Arg Tyr Ile  
 20 25 30  
 Ser Asp His Val Pro Val Gly Pro Arg Val Cys Val Gln Gly Ala Pro  
 35 40 45  
 Thr Ser Pro Ala Leu Cys Asn Ala Val Leu Leu Arg Leu Asp Arg Arg  
 50 55 60  
 Leu Ala Gly Leu Ala  
 65

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 69 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr  
 1 5 10 15  
 Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr  
 20 25 30  
 Ile Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys  
 35 40 45  
 Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro  
 50 55 60  
 Phe Arg Lys Gln Asn  
 65

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 69 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: not relevant  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Val Leu Pro Glu Leu Tyr Phe Met Lys Phe Asp Val Lys Ser Cys Tyr
1           5           10           15
Asp Ser Ile Pro Arg Met Glu Cys Met Arg Ile Leu Lys Asp Ala Leu
          20           25           30
Lys Asn Lys Cys Tyr Ile Arg Glu Asp Gly Leu Phe Gln Gly Ser Ser
          35           40           45
Leu Ser Ala Pro Ile Val Asp Leu Val Tyr Asp Asp Leu Leu Glu Phe
          50           55           60
Tyr Ser Glu Phe Lys
65
```

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 54 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: not relevant  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu Ile Thr Thr Gln Glu Asn
1           5           10           15
Asn Ala Val Leu Phe Ile Glu Lys Leu Ile Asn Val Ser Arg Glu Asn
          20           25           30
Gly Phe Lys Phe Asn Met Lys Lys Leu Gln Thr Gln Asp Tyr Cys Asp
          35           40           45
Trp Ile Gly Ile Ser Ile
50
```

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 54 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: not relevant  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

His Leu Ile Tyr Met Asp Asp Ile Lys Leu Tyr Ala Lys Asn Asp Lys  
1 5 10 15  
Glu Met Lys Lys Leu Ile Asp Thr Thr Thr Ile Phe Ser Asn Asp Ile  
20 25 30  
Ser Met Gln Phe Gly Leu Asp Lys Cys Lys Thr Lys Cys Leu Tyr Lys  
35 40 45  
Tyr Leu Gly Phe Gln Gln  
50

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Tyr Val Arg Tyr Ala Asp Asp Ile Leu Ile Gly Val Leu Gly Ser Lys  
1 5 10 15  
Asn Lys Ile Ile Lys Arg Asp Leu Asn Asn Phe Leu Asn Ser Leu Gly  
20 25 30  
Leu Thr Ile Asn Glu Glu Lys Thr Leu Ile Glu Thr Pro Ala Arg Phe  
35 40 45  
Leu Gly Tyr Asn Ile  
50

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser His Leu Glu Ile  
1 5 10 15  
Gly His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp  
20 25 30  
Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu  
35 40 45  
Trp Met Gly Tyr Glu Leu  
50



(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 49 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: not relevant  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Ile Leu Lys Leu Ala Asp Asp Phe Leu Ile Ile Ser Thr Asp Gln Gln
1           5           10           15
Gln Val Ile Asn Ile Lys Lys Leu Ala Met Gly Gly Phe Gln Lys Tyr
          20           25           30
Asn Ala Lys Ala Asn Arg Ile Arg Ser Lys Ser Ser Lys Gly Ile Phe
          35           40           45
```

Arg

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 39 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: not relevant  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Leu Gln Lys Gln Leu Glu Phe Tyr Phe Ser Asp Ala Asn Leu Tyr Asn
1           5           10           15
Asp Ser Phe Leu Arg Lys Leu Val Leu Lys Ser Gly Glu Gln Arg Val
          20           25           30
Glu Ile Glu Thr Leu Leu Met
          35
```

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 37 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: not relevant  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ile Cys His Gln Glu Tyr Tyr Phe Gly Asp Phe Asn Leu Pro Arg Asp  
1 5 10 15  
Lys Phe Leu Lys Glu Gln Ile Lys Leu Asp Glu Gly Trp Val Pro Leu  
20 25 30  
Glu Ile Met Ile Lys  
35

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ile Cys Glu Gln Ile Glu Tyr Tyr Phe Gly Asp His Asn Leu Pro Arg  
1 5 10 15  
Asp Lys Phe Leu Lys Gln Gln Ile Leu Leu Asp Asp Gly Trp Val Pro  
20 25 30  
Leu Glu Thr Met Ile Lys  
35

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ile Leu Arg Gln Val Glu Tyr Tyr Phe Gly Asp Ala Asn Leu Asn Arg  
1 5 10 15  
Asp Lys Phe Leu Arg Glu Gln Ile Gly Lys Asn Glu Asp Gly Trp Val  
20 25 30  
Pro Leu Ser Val Leu Val Thr  
35

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Cys Leu Lys Gln Val Glu Phe Tyr Phe Ser Glu Phe Asn Phe Pro Tyr  
1 5 10 15  
Asp Arg Phe Leu Arg Thr Thr Ala Glu Lys Asn Asp Gly Trp Val Pro  
20 25 30  
Ile Ser Thr Ile Ala Thr  
35

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TAGACCTGTT AGTGACATT TGAATTGAAG C 31

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TAGACCTGTT AGGTTGGATT TGTGGCATCA 30

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CAAAACCCCA AAACCTAACA GGTCTA 26

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 103 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCGGGAATTC TAATACGACT CACTATAGGG AAGAACTCT GATGAGGCCG AAAGGCCGAA 60  
ACTCCACGAA AGTGGAGTAA GTTTCTCGAT AATTGATCTG TAG 103

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CGGGGATCCT CTTCAAAAGA TGAGAGGACA GCAAAC 36

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CCCCAAAACC CAAAACCCC AAAACCCCCA CAGGGGTTTT GGGGTTTTGG GGTTTTGGGG 60

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 58 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CCAAAACCCC AAAACCCCAA AACCCCCACA GGGGTTTTGG GGTTTTGGGG TTTTGGGG 58

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 56 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AAAACCCCAA AACCCCAAAA CCCCACAGG GGTTTTGGGG TTTTGGGGTT TTGGGG

56

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 54 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AACCCCAAAA CCCCAAACC CCCACAGGGG TTTTGGGGTT TTGGGGTTTT GGGG

54

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CCCCAAACC CCAAACCCC CACAGGGGTT TTGGGGTTTT GGGGTTTT

48

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 52 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AAAACCCCAA AACCCCAAAA CCCCACAGG GGTTTTGGGG TTTTGGGGTT TT

52

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 50 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AACCCCAAAA CCCCAAACC CCCACAGGGG TTTTGGGGTT TTGGGGTTTT

50

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCCAAAACC CCAAACCCC CACAGGGGTT TTGGGGTTTT GGGGTTTT

48

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 46 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCAAACCCC AAAACCCCCA CAGGGGTTTT GGGGTTTTGG GGT

46

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 44 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAAACCCCAA AACCCCAACA GGGGTTTTGG GGT

44

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "RNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CAAAACCCCA AAACC

15

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TTTTG GGG

8

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "RNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CAAAACCCCA AAACC

15

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGGGTTTT

8

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCTRAARTAR TGDGTNADRT TRTTCAT

27

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GCGGATCCAT GAAYCCWGAR AAYCCWAAYG T

31

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

NNNGTNACHG GHATHAAYAA

20

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

DGCDGTYTCY TGRTCRTRT A

21



(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2421 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

AACTCATTTA ATTACTAATT TAATCAACAA GATTGATAAA AAGCAGTAAA TAAAACCCAA	60
TAGATTTAAT TTAGAAAAGTA TCAATTGAAA AATGGAAAATT GAAAACAACT AAGCACAATA	120
GCCAAAAGCC GAAAAATTGT GGTGGGAACT TGAATTAGAG ATGCAAGAAA ACCAAAATGA	180
TATATAAGTT AGGGTTAAGA TTGACGATCC TAAGCAATAT CTCGTGAACG TCACTGCAGC	240
ATGTTTGTG TAGGAAGGTA GTTACTACTA AGATAAAGAT GAAAGAAGAT ATATCATCAC	300
TAAAGCACTT CTTGAGGTGG CTGAGTCTGA TCCTGAGTTC ATCTGCTAGT TGGCAGTCTA	360
CATCCGTAAT GAACTTTACA TCAGAACTAC CACTAACTAC ATTGTAGCAT TTTGTGTTGT	420
CCACAAGAAT ACTCAACCAT TCATCGAAAA GTACTTCAAC AAAGCAGTAC TTTTGCCTAA	480
TGACTTACTG GAAGTCTGTG AATTTGCATA GGTCTCTAT ATTTTGTATG CAACTGAATT	540
CAAAAATTTG TATCTTGATA GGATACTTTC ATAAGATATT CGTAAGGAAC TCACTTTCCG	600
TAAGTGTTTA CAAAGATGCG TCAGAAGCAA GTTTTCTGAA TTCAACGAAT ACTAAGTTGG	660
TAAGTATTGC ACTGAATCCT AACGTAAGAA AACAATGTTC CGTTACCTCT CAGTTACCAA	720
CAAGTAAAAG TGGGATTAAA CTAAGAAGAA GAGAAAAGAG AATCTCTTAA CCAAACTTTA	780
GGCAATAAAG GAATCTGAAG ATAAGTCCAA GAGAGAACT GGAGACATAA TGAACGTTGA	840
AGATGCAATC AAGGCTTTAA AACCAGCAGT TATGAAGAAA ATAGCCAAGA GATAGAATGC	900
CATGAAGAAA CACATGAAGG CACCTAAAAT TCCTAACTCT ACCTTGGAAT CAAAGTACTT	960
GACCTTCAAG GATCTCATT AGTTC TGCCA TATTTCTGAG CCTAAAGAAA GAGTCTATAA	1020
GATCCTTGGT AAAAAATACC CTAAGACCGA AGAGGAATAC AAAGCAGCCT TTGGTGATTC	1080
TGCATCTGCA CCCTTCAATC CTGAATTGGC TGGAAAGCGT ATGAAGATTG AAATCTCTAA	1140
AACATGGGAA AATGAACTCA GTGCAAAAGG CAACACTGCT GAGGTTTGGG ATAATTTAAT	1200
TTCAAGCAAT TAACTCCCAT ATATGGCCAT GTTACGTAAC TTGTCTAACA TCTTAAAAGC	1260
CGGTGTTTCA GATACTACAC ACTCTATTGT GATCAACAAG ATTTGTGAGC CCAAGGCCGT	1320
TGAGAACTCC AAGATGTTCC CTCTTCAATT CTTTAGTGCC ATTGAAGCTG TTAATGAAGC	1380
AGTTACTAAG GGATTCAAGG CCAAGAAGAG AGAAAATATG AATCTTAAAG GTCAAATCGA	1440
AGCAGTAAAG GAAGTTGTTG AAAAAACCGA TGAAGAGAAG AAAGATATGG AGTTGGAGTA	1500

AACCGAAGAA GGAGAATTTG TTAAAGTCAA CGAAGGAATT GGCAAGCAAT ACATTAAGTC	1560
CATTGAACTT GCAATCAAGA TAGCAGTTAA CAAGAATTTA GATGAAATCA AAGGACACAC	1620
TGCAATCTTC TCTGATGTTT CTGGTTCTAT GAGTACCTCA ATGTCAGGTG GAGCCAAGAA	1680
GTATGGTTCC GTTCGTACTT GTCTCGAGTG TGCATTAGTC CTTGGTTTGA TGGTAAAATA	1740
ACGTTGTGAA AAGTCCTCAT TCTACATCTT CAGTTCACCT AGTTCTCAAT GCAATAAGTG	1800
TTACTTAGAA GTTGATCTCC CTGGAGACGA ACTCCGTCCT TCTATGTAAA AACTTTTGCA	1860
AGAGAAAGGA AAACCTGGTG GTGGTACTGA TTTCCCTAT GAGTGCATTG ATGAATGGAC	1920
AAAGAATAAA ACTCACGTAG ACAATATCGT TATTTTGTCT GATATGATGA TTGCAGAAGG	1980
ATATTCAGAT ATCAATGTTA GAGGCAGTTC CATTGTAAAC AGCATCAAAA AGTACAAGGA	2040
TGAAGTAAAT CCTAACATTA AAATCTTTGC AGTTGACTTA GAAGGTTACG GAAAGTGCCT	2100
TAATCTAGGT GATGAGTTCA ATGAAAACAA CTACATCAAG ATATTCGGTA TGAGCGATTC	2160
AATCTTAAAG TTCATTTTCAG CCAAGCAAGG AGGAGCAAAT ATGGTCGAAG TTATCAAAAA	2220
CTTTGCCCTT CAAAAATAG GACAAAAGTG AGTTTCTTGA GATTCTTCTA TAACAAAAAT	2280
CTCACCCAC TTTTTTGTTC TATTGCATAG CCATTATGAA ATTTAAATTA TTATCTATTT	2340
ATTTAAGTTA CTTACATAGT TTATGTATCG CAGTCTATTA GCCTATTCAA ATGATTCTGC	2400
AAAGAACAAA AAAGATTAAA A	2421

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 699 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Glu	Leu	Glu	Leu	Glu	Met	Gln	Glu	Asn	Gln	Asn	Asp	Ile	Gln	Val	Arg
1				5					10					15	
Val	Lys	Ile	Asp	Asp	Pro	Lys	Gln	Tyr	Leu	Val	Asn	Val	Thr	Ala	Ala
			20					25					30		
Cys	Leu	Leu	Gln	Glu	Gly	Ser	Tyr	Tyr	Gln	Asp	Lys	Asp	Glu	Arg	Arg
			35				40					45			
Tyr	Ile	Ile	Thr	Lys	Ala	Leu	Glu	Val	Ala	Glu	Ser	Asp	Pro	Glu	
			50			55				60					
Phe	Ile	Cys	Gln	Leu	Ala	Val	Tyr	Ile	Arg	Asn	Glu	Leu	Tyr	Ile	Arg
65				70					75					80	
Thr	Thr	Thr	Asn	Tyr	Ile	Val	Ala	Phe	Cys	Val	Val	His	Lys	Asn	Thr
			85					90						95	

Gln Pro Phe Ile Glu Lys Tyr Phe Asn Lys Ala Val Leu Leu Pro Asn  
 100 105 110  
 Asp Leu Leu Glu Val Cys Glu Phe Ala Gln Val Leu Tyr Ile Phe Asp  
 115 120 125  
 Ala Thr Glu Phe Lys Asn Leu Tyr Leu Asp Arg Ile Leu Ser Gln Asp  
 130 135 140  
 Ile Arg Lys Glu Leu Thr Phe Arg Lys Cys Leu Gln Arg Cys Val Arg  
 145 150 155 160  
 Ser Lys Phe Ser Glu Phe Asn Glu Tyr Gln Leu Gly Lys Tyr Cys Thr  
 165 170 175  
 Glu Ser Gln Arg Lys Lys Thr Met Phe Arg Tyr Leu Ser Val Thr Asn  
 180 185 190  
 Lys Gln Lys Trp Asp Gln Thr Lys Lys Lys Arg Lys Glu Asn Leu Leu  
 195 200 205  
 Thr Lys Leu Gln Ala Ile Lys Glu Ser Glu Asp Lys Ser Lys Arg Glu  
 210 215 220  
 Thr Gly Asp Ile Met Asn Val Glu Asp Ala Ile Lys Ala Leu Lys Pro  
 225 230 235 240  
 Ala Val Met Lys Lys Ile Ala Lys Arg Gln Asn Ala Met Lys Lys His  
 245 250 255  
 Met Lys Ala Pro Lys Ile Pro Asn Ser Thr Leu Glu Ser Lys Tyr Leu  
 260 265 270  
 Thr Phe Lys Asp Leu Ile Lys Phe Cys His Ile Ser Glu Pro Lys Glu  
 275 280 285  
 Arg Val Tyr Lys Ile Leu Gly Lys Lys Tyr Pro Lys Thr Glu Glu Glu  
 290 295 300  
 Tyr Lys Ala Ala Phe Gly Asp Ser Ala Ser Ala Pro Phe Asn Pro Glu  
 305 310 315 320  
 Leu Ala Gly Lys Arg Met Lys Ile Glu Ile Ser Lys Thr Trp Glu Asn  
 325 330 335  
 Glu Leu Ser Ala Lys Gly Asn Thr Ala Glu Val Trp Asp Asn Leu Ile  
 340 345 350  
 Ser Ser Asn Gln Leu Pro Tyr Met Ala Met Leu Arg Asn Leu Ser Asn  
 355 360 365  
 Ile Leu Lys Ala Gly Val Ser Asp Thr Thr His Ser Ile Val Ile Asn  
 370 375 380  
 Lys Ile Cys Glu Pro Lys Ala Val Glu Asn Ser Lys Met Phe Pro Leu  
 385 390 395 400

Gln Phe Phe Ser Ala Ile Glu Ala Val Asn Glu Ala Val Thr Lys Gly  
 405 410 415  
 Phe Lys Ala Lys Lys Arg Glu Asn Met Asn Leu Lys Gly Gln Ile Glu  
 420 425 430  
 Ala Val Lys Glu Val Val Glu Lys Thr Asp Glu Glu Lys Lys Asp Met  
 435 440 445  
 Glu Leu Glu Gln Thr Glu Glu Gly Glu Phe Val Lys Val Asn Glu Gly  
 450 455 460  
 Ile Gly Lys Gln Tyr Ile Asn Ser Ile Glu Leu Ala Ile Lys Ile Ala  
 465 470 475 480  
 Val Asn Lys Asn Leu Asp Glu Ile Lys Gly His Thr Ala Ile Phe Ser  
 485 490 495  
 Asp Val Ser Gly Ser Met Ser Thr Ser Met Ser Gly Gly Ala Lys Lys  
 500 505 510  
 Tyr Gly Ser Val Arg Thr Cys Leu Glu Cys Ala Leu Val Leu Gly Leu  
 515 520 525  
 Met Val Lys Gln Arg Cys Glu Lys Ser Ser Phe Tyr Ile Phe Ser Ser  
 530 535 540  
 Pro Ser Ser Gln Cys Asn Lys Cys Tyr Leu Glu Val Asp Leu Pro Gly  
 545 550 555 560  
 Asp Glu Leu Arg Pro Ser Met Gln Lys Leu Leu Gln Glu Lys Gly Lys  
 565 570 575  
 Leu Gly Gly Gly Thr Asp Phe Pro Tyr Glu Cys Ile Asp Glu Trp Thr  
 580 585 590  
 Lys Asn Lys Thr His Val Asp Asn Ile Val Ile Leu Ser Asp Met Met  
 595 600 605  
 Ile Ala Glu Gly Tyr Ser Asp Ile Asn Val Arg Gly Ser Ser Ile Val  
 610 615 620  
 Asn Ser Ile Lys Lys Tyr Lys Asp Glu Val Asn Pro Asn Ile Lys Ile  
 625 630 635 640  
 Phe Ala Val Asp Leu Glu Gly Tyr Gly Lys Cys Leu Asn Leu Gly Asp  
 645 650 655  
 Glu Phe Asn Glu Asn Asn Tyr Ile Lys Ile Phe Gly Met Ser Asp Ser  
 660 665 670  
 Ile Leu Lys Phe Ile Ser Ala Lys Gln Gly Gly Ala Asn Met Val Glu  
 675 680 685  
 Val Ile Lys Asn Phe Ala Leu Gln Lys Ile Gly  
 690 695

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2829 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TCAATACTAT TAATTAATAA ATAAAAAAAA GCAAACCTACA AAGAAAATGT CAAGGCGTAA	60
CTAAAAAAAAAG CCATAGGCTC CTATAGGCAA TGAAACAAAT CTTGATTTTG TATTACAAAA	120
TCTAGAAGTT TACAAAAGCC AGATTGAGCA TTATAAGACC TAGTAGTAAT AGATCAAAGA	180
GGAGGATCTC AAGCTTTTAA AGTTCAAAAA TTAAGATTAG GATGGAAACT CTGGCAACGA	240
TGATGATGAT GAAGAAAACA ACTCAAATAA ATAATAAGAA TTATTAAGGA GAGTCAATTA	300
GATTAAGTAG CAAGTTTAAT TGATAAAAAA AGTTGGTTCT AAGGTAGAGA AAGATTTGAA	360
TTTGAACGAA GATGAAAACA AAAAGAATGG ACTTTCTGAA TAGCAAGTGA AAGAAGAGTA	420
ATTAAGAACG ATTACTGAAG AATAGGTTAA GTATTAAAAT TTAGTATTTA ACATGGACTA	480
CCAGTTAGAT TTAAATGAGA GTGGTGGCCA TAGAAGACAC AGAAGAGAAA CAGATTATGA	540
TACTGAAAAA TGGTTTGAAA TATCTCATGA CCAAAAAAAT TATGTATCAA TTTACGCCAA	600
CTAAAAGACA TCATATTGTT GGTGGCTTAA AGATTATTTT AATAAAAAACA ATTATGATCA	660
TCTTAATGTA AGCATTAAAC GACTAGAAAC TGAAGCCGAA TTCTATGCCT TTGATGATTT	720
TTCAACAAAC ATCAAACTTA CTAATAATTC TTACTAGACT GTTAACATAG ACGTTAATTT	780
TGATAATAAT CTCTGTATAC TCGCATTGCT TAGATTTTTT TTATCACTAG AAAGATTCAA	840
TATTTTGAAT ATAAGATCTT CTTATACAAG AAATTAATAT AATTTTGAGA AAATTGGTGA	900
GCTACTTGAA ACTATCTTCG CAGTTGTCTT TTCTCATCGC CACTTACAAG GCATTCATTT	960
ACAAGTTCCT TGCGAAGCGT TCTAATATTT AGTTAACTCC TCATCATAAA TTAGCGTTAA	1020
AGATAGCTAA TTATAGGTAT ACTCTTTCTC TACAGACTTA AAATTAGTTG AACTAACAA	1080
AGTCCAAGAT TATTTTAAGT TCTTATAAGA ATTCCCTCGT TTGACTCATG TAAGCTAGTA	1140
GGCTATCCCA GTTAGTGCTA CTAACGCTGT AGAGAACCTC AATGTTTTAC TTAAAAAGGT	1200
CAAGCATGCT AATCTTAATT TAGTTTCTAT CCCTACCTAA TTCAATTTTG ATTTCTACTT	1260
TGTTAATTTA TAACATTTGA AATTAGAGTT TGGATTAGAA CCAAATATTT TGACAAAACA	1320
AAAGCTTGAA AATCTACTTT TGAGTATAAA ATAATCAAAA AATCTTAAAT TTTTAAGATT	1380
AACTTTTAC ACCTACGTTG CTTAAGAAAC CTCCAGAAAA CAGATATTAA AACAAGCTAC	1440
AACAATCAAA AATCTCAAAA ACAATAAAAA TCAAGAAGAA ACTCCTGAAA CTAAAGATGA	1500
AACTCCAAGC GAAAGCACAA GTGGTATGAA ATTTTTTGAT CATCTTTCTG AATTAACCGA	1560
GCTTGAAGAT TTCAGCGTTA ACTTGTAAGC TACCCAAGAA ATTTATGATA GCTTGCACAA	1620

ACTTTTGATT AGATCAACAA ATTTAAAGAA GTTCAAATTA AGTTACAAAT ATGAAATGGA	1680
AAAGAGTAAA ATGGATACAT TCATAGATCT TAAGAATATT TATGAAACCT TAAACAATCT	1740
TAAAAGATGC TCTGTTAATA TATCAAATCC TCATGGAAAC ATTTCTTATG AACTGACAAA	1800
TAAAGATTCT ACTTTTTATA AATTTAAGCT GACCTTAAAC TAAGAATTAT AACACGCTAA	1860
GTATACTTTT AAGTAGAACG AATTTTAATT TAATAACGTT AAAAGTGCAA AAATTGAATC	1920
TTCCTCATTG GAAAGCTTAG AAGATATTGA TAGTCTTTGC AAATCTATTG CTTCTTGTA	1980
AAATTTACAA AATGTTAATA TTATCGCCAG TTTGCTCTAT CCCAACAATA TTTAGAAAAA	2040
TCCTTTCAAT AAGCCCAATC TTCTATTTTT CAAGCAATTT GAATAATTGA AAAATTTGGA	2100
AAATGTATCT ATCAACTGTA TTCTTGATCA GCATATACTT AATTCTATTT CAGAATTCCT	2160
AGAAAAGAAT AAAAAAATAA AAGCATTTCAT TTTGAAAAGA TATTATTTAT TACAATATTA	2220
TCTTGATTAT ACTAAATTAT TTAAACACT TCAATAGTTA CCTGAATTAA ATTAAGTTTA	2280
CATTAATTAG CAATTAGAAG AATTGACTGT GAGTGAAGTA CATAAGTAAG TATGGGAAAA	2340
CCACAAGCAA AAAGCTTTCT ATGAACCATT ATGTGAGTTT ATCAAAGAAT CATCCTAAAC	2400
CCTTTAGCTA ATAGATTTTG ACCAAAACAC TGTAAGTGAT GACTCTATTA AAAAGATTTT	2460
AGAATCTATA TCTGAGTCTA AGTATCATCA TTATTTGAGA TTGAACCCTA GTTAATCTAG	2520
CAGTTTAATT AAATCTGAAA ACGAAGAAAT TTAAGAACTT CTCAAAGCTT GCGACGAAAA	2580
AGGTGTTTTA GTAAAAGCAT ACTATAAATT CCCTCTATGT TTACCAACTG GTACTTATTA	2640
CGATTACAAT TCAGATAGAT GGTGATTAAT TAAATATTAG TTTAAATAAA TATTAAATAT	2700
TGAATATTTT TTTGCTTATT ATTTGAATAA TACATACAAT AGTCATTTTT AGTGTTTTGA	2760
ATATATTTTA GTTATTTAAT TCATTATTTT AAGTAAATAA TTATTTTTC AATCATTTTT	2820
AAAAAATCG	2829

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 872 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Met	Ser	Arg	Arg	Asn	Gln	Lys	Lys	Pro	Gln	Ala	Pro	Ile	Gly	Asn	Glu
1				5				10					15		
Thr	Asn	Leu	Asp	Phe	Val	Leu	Gln	Asn	Leu	Glu	Val	Tyr	Lys	Ser	Gln
		20					25					30			
Ile	Glu	His	Tyr	Lys	Thr	Gln	Gln	Gln	Gln	Ile	Lys	Glu	Glu	Asp	Leu
		35				40						45			

Lys Leu Leu Lys Phe Lys Asn Gln Asp Gln Asp Gly Asn Ser Gly Asn  
 50 55 60  
 Asp Asp Asp Asp Glu Glu Asn Asn Ser Asn Lys Gln Gln Glu Leu Leu  
 65 70 75 80  
 Arg Arg Val Asn Gln Ile Lys Gln Gln Val Gln Leu Ile Lys Lys Val  
 85 90 95  
 Gly Ser Lys Val Glu Lys Asp Leu Asn Leu Asn Glu Asp Glu Asn Lys  
 100 105 110  
 Lys Asn Gly Leu Ser Glu Gln Gln Val Lys Glu Glu Gln Leu Arg Thr  
 115 120 125  
 Ile Thr Glu Glu Gln Val Lys Tyr Gln Asn Leu Val Phe Asn Met Asp  
 130 135 140  
 Tyr Gln Leu Asp Leu Asn Glu Ser Gly Gly His Arg Arg His Arg Arg  
 145 150 155 160  
 Glu Thr Asp Tyr Asp Thr Glu Lys Trp Phe Glu Ile Ser His Asp Gln  
 165 170 175  
 Lys Asn Tyr Val Ser Ile Tyr Ala Asn Gln Lys Thr Ser Tyr Cys Trp  
 180 185 190  
 Trp Leu Lys Asp Tyr Phe Asn Lys Asn Asn Tyr Asp His Leu Asn Val  
 195 200 205  
 Ser Ile Asn Arg Leu Glu Thr Glu Ala Glu Phe Tyr Ala Phe Asp Asp  
 210 215 220  
 Phe Ser Gln Thr Ile Lys Leu Thr Asn Asn Ser Tyr Gln Thr Val Asn  
 225 230 235 240  
 Ile Asp Val Asn Phe Asp Asn Asn Leu Cys Ile Leu Ala Leu Leu Arg  
 245 250 255  
 Phe Leu Leu Ser Leu Glu Arg Phe Asn Ile Leu Asn Ile Arg Ser Ser  
 260 265 270  
 Tyr Thr Arg Asn Gln Tyr Asn Phe Glu Lys Ile Gly Glu Leu Leu Glu  
 275 280 285  
 Thr Ile Phe Ala Val Val Phe Ser His Arg His Leu Gln Gly Ile His  
 290 295 300  
 Leu Gln Val Pro Cys Glu Ala Phe Gln Tyr Leu Val Asn Ser Ser Ser  
 305 310 315 320  
 Gln Ile Ser Val Lys Asp Ser Gln Leu Gln Val Tyr Ser Phe Ser Thr  
 325 330 335  
 Asp Leu Lys Leu Val Asp Thr Asn Lys Val Gln Asp Tyr Phe Lys Phe  
 340 345 350  
 Leu Gln Glu Phe Pro Arg Leu Thr His Val Ser Gln Gln Ala Ile Pro  
 355 360 365  
 Val Ser Ala Thr Asn Ala Val Glu Asn Leu Asn Val Leu Leu Lys Lys  
 370 375 380

Val Lys His Ala Asn Leu Asn Leu Val Ser Ile Pro Thr Gln Phe Asn  
385 390 395 400  
Phe Asp Phe Tyr Phe Val Asn Leu Gln His Leu Lys Leu Glu Phe Gly  
405 410 415  
Leu Glu Pro Asn Ile Leu Thr Lys Gln Lys Leu Glu Asn Leu Leu Leu  
420 425 430  
Ser Ile Lys Gln Ser Lys Asn Leu Lys Phe Leu Arg Leu Asn Phe Tyr  
435 440 445  
Thr Tyr Val Ala Gln Glu Thr Ser Arg Lys Gln Ile Leu Lys Gln Ala  
450 455 460  
Thr Thr Ile Lys Asn Leu Lys Asn Asn Lys Asn Gln Glu Glu Thr Pro  
465 470 475 480  
Glu Thr Lys Asp Glu Thr Pro Ser Glu Ser Thr Ser Gly Met Lys Phe  
485 490 495  
Phe Asp His Leu Ser Glu Leu Thr Glu Leu Glu Asp Phe Ser Val Asn  
500 505 510  
Leu Gln Ala Thr Gln Glu Ile Tyr Asp Ser Leu His Lys Leu Leu Ile  
515 520 525  
Arg Ser Thr Asn Leu Lys Lys Phe Lys Leu Ser Tyr Lys Tyr Glu Met  
530 535 540  
Glu Lys Ser Lys Met Asp Thr Phe Ile Asp Leu Lys Asn Ile Tyr Glu  
545 550 555 560  
Thr Leu Asn Asn Leu Lys Arg Cys Ser Val Asn Ile Ser Asn Pro His  
565 570 575  
Gly Asn Ile Ser Tyr Glu Leu Thr Asn Lys Asp Ser Thr Phe Tyr Lys  
580 585 590  
Phe Lys Leu Thr Leu Asn Gln Glu Leu Gln His Ala Lys Tyr Thr Phe  
595 600 605  
Lys Gln Asn Glu Phe Gln Phe Asn Asn Val Lys Ser Ala Lys Ile Glu  
610 615 620  
Ser Ser Ser Leu Glu Ser Leu Glu Asp Ile Asp Ser Leu Cys Lys Ser  
625 630 635 640  
Ile Ala Ser Cys Lys Asn Leu Gln Asn Val Asn Ile Ile Ala Ser Leu  
645 650 655  
Leu Tyr Pro Asn Asn Ile Gln Lys Asn Pro Phe Asn Lys Pro Asn Leu  
660 665 670  
Leu Phe Phe Lys Gln Phe Glu Gln Leu Lys Asn Leu Glu Asn Val Ser  
675 680 685  
Ile Asn Cys Ile Leu Asp Gln His Ile Leu Asn Ser Ile Ser Glu Phe  
690 695 700  
Leu Glu Lys Asn Lys Lys Ile Lys Ala Phe Ile Leu Lys Arg Tyr Tyr  
705 710 715 720





Ala Ala Trp Lys Met Phe His Ser Leu Val Gly Thr Tyr Ala Phe Val  
 115 120 125  
 Asp Leu Leu Ile Asn Tyr Thr Val Ile Gln Phe Asn Gly Gln Phe Phe  
 130 135 140  
 Thr Gln Ile Val Gly Asn Arg Cys Asn Glu Pro His Leu Pro Pro Lys  
 145 150 155 160  
 Trp Val Gln Arg Ser Ser Ser Ser Ser Ala Thr Ala Ala Gln Ile Lys  
 165 170 175  
 Gln Leu Thr Glu Pro Val Thr Asn Lys Gln Phe Leu His Lys Leu Asn  
 180 185 190  
 Ile Asn Ser Ser Ser Phe Phe Pro Tyr Ser Lys Ile Leu Pro Ser Ser  
 195 200 205  
 Ser Ser Ile Lys Lys Leu Thr Asp Leu Arg Glu Ala Ile Phe Pro Thr  
 210 215 220  
 Asn Leu Val Lys Ile Pro Gln Arg Leu Lys Val Arg Ile Asn Leu Thr  
 225 230 235 240  
 Leu Gln Lys Leu Leu Lys Arg His Lys Arg Leu Asn Tyr Val Ser Ile  
 245 250 255  
 Leu Asn Ser Ile Cys Pro Pro Leu Glu Gly Thr Val Leu Asp Leu Ser  
 260 265 270  
 His Leu Ser Arg Gln Ser Pro Lys Glu Arg Val Leu Lys Phe Ile Ile  
 275 280 285  
 Val Ile Leu Gln Lys Leu Leu Pro Gln Glu Met Phe Gly Ser Lys Lys  
 290 295 300  
 Asn Lys Gly Lys Ile Ile Lys Asn Leu Asn Leu Leu Leu Ser Leu Pro  
 305 310 315 320  
 Leu Asn Gly Tyr Leu Pro Phe Asp Ser Leu Leu Lys Lys Leu Arg Leu  
 325 330 335  
 Lys Asp Phe Arg Trp Leu Phe Ile Ser Asp Ile Trp Phe Thr Lys His  
 340 345 350  
 Asn Phe Glu Asn Leu Asn Gln Leu Ala Ile Cys Phe Ile Ser Trp Leu  
 355 360 365  
 Phe Arg Gln Leu Ile Pro Lys Ile Ile Gln Thr Phe Phe Tyr Cys Thr  
 370 375 380  
 Glu Ile Ser Ser Thr Val Thr Ile Val Tyr Phe Arg His Asp Thr Trp  
 385 390 395 400  
 Asn Lys Leu Ile Thr Pro Phe Ile Val Glu Tyr Phe Lys Thr Tyr Leu  
 405 410 415  
 Val Glu Asn Asn Val Cys Arg Asn His Asn Ser Tyr Thr Leu Ser Asn  
 420 425 430  
 Phe Asn His Ser Lys Met Arg Ile Ile Pro Lys Lys Ser Asn Asn Glu  
 435 440 445

Phe Arg Ile Ile Ala Ile Pro Cys Arg Gly Ala Asp Glu Glu Glu Phe  
450 455 460  
Thr Ile Tyr Lys Glu Asn His Lys Asn Ala Ile Gln Pro Thr Gln Lys  
465 470 475 480  
Ile Leu Glu Tyr Leu Arg Asn Lys Arg Pro Thr Ser Phe Thr Lys Ile  
485 490 495  
Tyr Ser Pro Thr Gln Ile Ala Asp Arg Ile Lys Glu Phe Lys Gln Arg  
500 505 510  
Leu Leu Lys Lys Phe Asn Asn Val Leu Pro Glu Leu Tyr Phe Met Lys  
515 520 525  
Phe Asp Val Lys Ser Cys Tyr Asp Ser Ile Pro Arg Met Glu Cys Met  
530 535 540  
Arg Ile Leu Lys Asp Ala Leu Lys Asn Glu Asn Gly Phe Phe Val Arg  
545 550 555 560  
Ser Gln Tyr Phe Phe Asn Thr Asn Thr Gly Val Leu Lys Leu Phe Asn  
565 570 575  
Val Val Asn Ala Ser Arg Val Pro Lys Pro Tyr Glu Leu Tyr Ile Asp  
580 585 590  
Asn Val Arg Thr Val His Leu Ser Asn Gln Asp Val Ile Asn Val Val  
595 600 605  
Glu Met Glu Ile Phe Lys Thr Ala Leu Trp Val Glu Asp Lys Cys Tyr  
610 615 620  
Ile Arg Glu Asp Gly Leu Phe Gln Gly Ser Ser Leu Ser Ala Pro Ile  
625 630 635 640  
Val Asp Leu Val Tyr Asp Asp Leu Leu Glu Phe Tyr Ser Glu Phe Lys  
645 650 655  
Ala Ser Pro Ser Gln Asp Thr Leu Ile Leu Lys Leu Ala Asp Asp Phe  
660 665 670  
Leu Ile Ile Ser Thr Asp Gln Gln Gln Val Ile Asn Ile Lys Lys Leu  
675 680 685  
Ala Met Gly Gly Phe Gln Lys Tyr Asn Ala Lys Ala Asn Arg Asp Lys  
690 695 700  
Ile Leu Ala Val Ser Ser Gln Ser Asp Asp Asp Thr Val Ile Gln Phe  
705 710 715 720  
Cys Ala Met His Ile Phe Val Lys Glu Leu Glu Val Trp Lys His Ser  
725 730 735  
Ser Thr Met Asn Asn Phe His Ile Arg Ser Lys Ser Ser Lys Gly Ile  
740 745 750  
Phe Arg Ser Leu Ile Ala Leu Phe Asn Thr Arg Ile Ser Tyr Lys Thr  
755 760 765  
Ile Asp Thr Asn Leu Asn Ser Thr Asn Thr Val Leu Met Gln Ile Asp  
770 775 780

His Val Val Lys Asn Ile Ser Glu Cys Tyr Lys Ser Ala Phe Lys Asp  
 785 790 795 800  
 Leu Ser Ile Asn Val Thr Gln Asn Met Gln Phe His Ser Phe Leu Gln  
 805 810 815  
 Arg Ile Ile Glu Met Thr Val Ser Gly Cys Pro Ile Thr Lys Cys Asp  
 820 825 830  
 Pro Leu Ile Glu Tyr Glu Val Arg Phe Thr Ile Leu Asn Gly Phe Leu  
 835 840 845  
 Glu Ser Leu Ser Ser Asn Thr Ser Lys Phe Lys Asp Asn Ile Ile Leu  
 850 855 860  
 Leu Arg Lys Glu Ile Gln His Leu Gln Ala Tyr Ile Tyr Ile Tyr Ile  
 865 870 875 880  
 His Ile Val Asn

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

YARACHAARG GHATYCCHYA RGG

23

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "DNA"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

DGTDATNARN ARRTARTCRT C

21

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Leu Cys Val Ser Tyr Ile Leu Ser Ser Phe Tyr Tyr Ala Asn Leu Glu  
 1 5 10 15

Glu Asn Ala Leu Gln Phe Leu Arg Lys Glu Ser Met Asp Pro Glu Lys  
 20 25 30

Pro Glu Thr Asn Leu Leu Met Arg Leu Thr  
 35 40

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Leu Cys Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr Leu Glu  
 1 5 10 15

Glu Ser Ser Leu Gly Phe Leu Arg Asp Glu Ser Met Asn Pro Glu Asn  
 20 25 30

Pro Asn Val Asn Leu Leu Met Arg Leu Thr  
 35 40

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "RNA"

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 12..25
- (D) OTHER INFORMATION: /mod\_base= OTHER

/note= "The residues located at these positions are  
 2-O-methylribonucleoti..."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TAGACCTGTT AGGUUUUGGG GUUUUG

26

(2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GGGGTTTTTG GGT

16

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 389 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..389
- (D) OTHER INFORMATION: /note= "expressed sequence tag (EST)  
AA281296"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GCCAAGTTCC TGC	ACTGGCT	GATGAGTGTG	TACGTCGTCG	AGCTGCTCAG	GTCTTTCTTT	60
TATGTCACGG	AGACCACGTT	TCAAAGAAC	AGGCTCTTTT	TCTACCGGAA	GAGTGTCTGG	120
AGCAAGTTGC	AAAGCATTGG	AATCAGACAG	CACTTGAAGA	GGGTGCAGCT	GCGGGACGTG	180
TCGGAAGCAG	AGGTCAGGCA	GCATCGGGAA	GCCAGGCCCG	CCCTGCTGAC	GTCCAGACTC	240
CGCTTCATCC	CCAAGCCTGA	CGGGCTGCGG	CCGATTGTGA	ACATGGACTA	CGTCGTGGGA	300
GCCAGAACGT	TCCGCAGAGA	AAAGAGGGCC	GAGCGTCTCA	CCTCGAGGGT	GAAGGCACTG	360
TTCAGCGTGC	TCAACTACGA	GCGGGCGCG				389

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..233
- (D) OTHER INFORMATION: /note= "TRT motifs from  
Schizosaccharomyces pombe tez1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Ile	Ser	Glu	Ile	Glu	Trp	Leu	Val	Leu	Gly	Lys	Arg	Ser	Asn	Ala	Lys
1				5					10					15	
Met	Cys	Leu	Ser	Asp	Phe	Glu	Lys	Arg	Lys	Gln	Ile	Phe	Ala	Glu	Phe
				20				25					30		

Ile Tyr Trp Leu Tyr Asn Ser Phe Ile Ile Pro Ile Leu Gln Ser Phe  
           35                                  40                                  45  
 Phe Tyr Ile Thr Glu Ser Ser Asp Leu Arg Asn Arg Thr Val Tyr Phe  
           50                                  55                                  60  
 Arg Lys Asp Ile Trp Lys Leu Leu Cys Arg Pro Phe Ile Thr Ser Met  
   65                                  70                                  75                                  80  
 Lys Met Glu Ala Phe Glu Lys Ile Asn Glu Asn Asn Val Arg Met Asp  
                                   85                                  90                                  95  
 Thr Gln Lys Thr Thr Leu Pro Pro Ala Val Ile Arg Leu Leu Pro Lys  
                                  100                                 105                                 110  
 Lys Asn Thr Phe Arg Leu Ile Thr Asn Leu Arg Lys Arg Phe Leu Ile  
                                  115                                 120                                 125  
 Lys Met Gly Ser Asn Lys Lys Met Leu Val Ser Thr Asn Gln Thr Leu  
   130                                 135                                 140  
 Arg Pro Val Ala Ser Ile Leu Lys His Leu Ile Asn Glu Glu Ser Ser  
  145                                 150                                 155                                 160  
 Gly Ile Pro Phe Asn Leu Glu Val Tyr Met Lys Leu Leu Thr Phe Lys  
                                  165                                 170                                 175  
 Lys Asp Leu Leu Lys His Arg Met Phe Gly Arg Lys Lys Tyr Phe Val  
                                  180                                 185                                 190  
 Arg Ile Asp Ile Lys Ser Cys Tyr Asp Arg Ile Lys Gln Asp Leu Met  
                                  195                                 200                                 205  
 Phe Arg Ile Val Lys Lys Lys Leu Lys Asp Pro Glu Phe Val Ile Arg  
   210                                 215                                 220  
 Lys Tyr Ala Thr Ile His Ala Thr Ser  
  225                                 230

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..233
- (D) OTHER INFORMATION: /note= "TRT motifs from *Saccharomyces cerevisiae* EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Leu Lys Asp Phe Arg Trp Leu Phe Ile Ser Asp Ile Trp Phe Thr Lys  
 1                                  5                                  10                                  15

His Asn Phe Glu Asn Leu Asn Gln Leu Ala Ile Cys Phe Ile Ser Trp  
                   20                                  25                                  30  
 Leu Phe Arg Gln Leu Ile Pro Lys Ile Ile Gln Thr Phe Phe Tyr Cys  
                   35                                  40                                  45  
 Thr Glu Ile Ser Ser Thr Val Thr Ile Val Tyr Phe Arg His Asp Thr  
                   50                                  55                                  60  
 Trp Asn Lys Leu Ile Thr Pro Phe Ile Val Glu Tyr Phe Lys Thr Tyr  
                   65                                  70                                  75                                  80  
 Leu Val Glu Asn Asn Val Cys Arg Asn His Asn Ser Tyr Thr Leu Ser  
                                   85                                  90                                  95  
 Asn Phe Asn His Ser Lys Met Arg Ile Ile Pro Lys Lys Ser Asn Asn  
                                   100                                  105                                  110  
 Glu Phe Arg Ile Ile Ala Ile Pro Cys Arg Gly Ala Asp Glu Glu Glu  
                   115                                  120                                  125  
 Phe Thr Ile Tyr Lys Glu Asn His Lys Asn Ala Ile Gln Pro Thr Gln  
                   130                                  135                                  140  
 Lys Ile Leu Glu Tyr Leu Arg Asn Lys Arg Pro Thr Ser Phe Thr Lys  
                   145                                  150                                  155                                  160  
 Ile Tyr Ser Pro Thr Gln Ile Ala Asp Arg Ile Lys Glu Phe Lys Gln  
                                   165                                  170                                  175  
 Arg Leu Leu Lys Lys Phe Asn Asn Val Leu Pro Glu Leu Tyr Phe Met  
                   180                                  185                                  190  
 Lys Phe Asp Val Lys Ser Cys Tyr Asp Ser Ile Pro Arg Met Glu Cys  
                   195                                  200                                  205  
 Met Arg Ile Leu Lys Asp Ala Leu Lys Asn Glu Asn Gly Phe Phe Val  
                   210                                  215                                  220  
 Arg Ser Gln Tyr Phe Phe Asn Thr Asn  
                   225                                  230

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..233
- (D) OTHER INFORMATION: /note= "TRT motifs from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:



Thr Arg Glu Ile Ser Trp Met Gln Val Glu Thr Ser Ala Lys His Phe  
 1 5 10 15  
 Tyr Tyr Phe Asp His Glu Asn Ile Tyr Val Leu Trp Lys Leu Leu Arg  
 20 25 30  
 Trp Ile Phe Glu Asp Leu Val Val Ser Leu Ile Arg Cys Phe Phe Tyr  
 35 40 45  
 Val Thr Glu Gln Gln Lys Ser Tyr Ser Lys Thr Tyr Tyr Tyr Arg Lys  
 50 55 60  
 Asn Ile Trp Asp Val Ile Met Lys Met Ser Ile Ala Asp Leu Lys Lys  
 65 70 75 80  
 Glu Thr Leu Ala Glu Val Gln Glu Lys Glu Val Glu Glu Trp Lys Lys  
 85 90 95  
 Ser Leu Gly Phe Ala Pro Gly Lys Leu Arg Leu Ile Pro Lys Lys Thr  
 100 105 110  
 Thr Phe Arg Pro Ile Met Thr Phe Asn Lys Lys Ile Val Asn Ser Asp  
 115 120 125  
 Arg Lys Thr Thr Lys Leu Thr Thr Asn Thr Lys Leu Leu Asn Ser His  
 130 135 140  
 Leu Met Leu Lys Thr Leu Lys Asn Arg Met Phe Lys Asp Pro Phe Gly  
 145 150 155 160  
 Phe Ala Val Phe Asn Tyr Asp Asp Val Met Lys Lys Tyr Glu Glu Phe  
 165 170 175  
 Val Cys Lys Trp Lys Gln Val Gly Gln Pro Lys Leu Phe Phe Ala Thr  
 180 185 190  
 Met Asp Ile Glu Lys Cys Tyr Asp Ser Val Asn Arg Glu Lys Leu Ser  
 195 200 205  
 Thr Phe Leu Lys Thr Thr Lys Leu Leu Ser Ser Asp Phe Trp Ile Met  
 210 215 220  
 Thr Ala Gln Ile Leu Lys Arg Lys Asn  
 225 230

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2631 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: -
- (B) LOCATION: 1..2631
- (D) OTHER INFORMATION: /note= "Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATTTATACTC ATGAAAATCT TATTCGAGTT CATTCAAGAC AAGCTTGACA TTGATCTACA	60
GACCAACAGT ACTTACAAAG AAAATTTAAA ATGTGGTCAC TTCAATGGCC TCGATGAAAT	120
TCTAACTACG TGTTTCGCAC TACCAAATTC AAGAAAAATA GCATTACCAT GCCTTCCTGG	180
TGACTTAAGC CACAAAGCAG TCATTGATCA CTGCATCATT TACCTGTTGA CGGGCGAATT	240
ATACAACAAC GTACTAACAT TTGGCTATAA AATAGCTAGA AATGAAGATG TCAACAATAG	300
TCTTTTTTGC CATTCTGCAA ATGTTAACGT TACGTTACTG AAAGGCGCTG CTTGGAAAAT	360
GTTCCACAGT TTGGTCGGTA CATACGCATT CGTTGATTTA TTGATCAATT ATACAGTAAT	420
TCAATTTAAT GGGCAGTTTT TCACTCAAAT CGTGGGTAAC AGATGTAACG AACCTCATCT	480
GCCGCCCAAA TGGGTCCAAC GATCATCCTC ATCATCCGCA ACTGCTGCGC AAATCAAACA	540
ACTTACAGAA CCAGTGACAA ATAAACAATT CTTACACAAG CTCAATATAA ATTCCTCTTC	600
TTTTTTTCCT TATAGCAAGA TCCTTCCTTC ATCATCATCT ATCAAAAAGC TAACTGACTT	660
GAGAGAAGCT ATTTTTCCCA CAAATTTGGT TAAAATTCCT CAGAGACTAA AGGTACGAAT	720
TAATTTGACG CTGCAAAAGC TATTAAAGAG ACATAAGCGT TTGAATTACG TTTCTATTTT	780
GAATAGTATT TGCCCAACCAT TGGAAGGGAC CGTATTGGAC TTGTCGCATT TGAGTAGGCA	840
ATCACCAAAG GAACGAGTCT TGAAATTTAT CATTGTTATT TTACAGAAGT TATTACCCCA	900
AGAAATGTTT GGCTCAAAGA AAAATAAAGG AAAAATTATC AAGAATCTAA ATCTTTTATT	960
AAGTTTACCC TTAAATGGCT ATTTACCATT TGATAGTTTG TTGAAAAAGT TAAGATTAAA	1020
GGATTTTCGG TGGTTGTTCA TTTCTGATAT TTGGTTCACC AAGCACAATT TTGAAAACCTT	1080
GAATCAATTG GCGATTTGTT TCATTTCTTG GCTATTTAGA CAACTAATTC CCAAAATTAT	1140
ACAGACTTTT TTTTACTGCA CCGAAATATC TTCTACAGTG ACAATTGTTT ACTTTAGACA	1200
TGATACTTGG AATAAACTTA TCACCCCTTT TATCGTAGAA TATTTTAAGA CGTACTTAGT	1260
CGAAAACAAC GTATGTAGAA ACCATAATAG TTACACGTTG TCCAATTTCA ATCATAGCAA	1320
AATGAGGATT ATACCAGAAA AAAGTAATAA TGAGTTCAGG ATTATTGCCA TCCCATGCAG	1380
AGGGGCAGAC GAAGAAGAAT TCACAATTTA TAAGGAGAAT CACAAAAATG CTATCCAGCC	1440
CACTCAAAAA ATTTTAGAAT ACCTAAGAAA CAAAAGGCCG ACTAGTTTFA CTAAAAATATA	1500
TTCTCCAACG CAAATAGCTG ACCGTATCAA AGAATTTAAG CAGAGACTTT TAAAGAAATT	1560
TAATAATGTC TTACCAGAGC TTTATTTTCAT GAAATTTGAT GTCAAATCTT GCTATGATTC	1620
CATACCAAGG ATGGAATGTA TGAGGATACT CAAGGATGCG CTAAAAAATG AAAATGGGTT	1680
TTTCGTTAGA TCTCAATATT TCTTCAATAC CAATACAGGT GTATTGAAGT TATTTAATGT	1740
TGTTAACGCT AGCAGAGTAC CAAAACCTTA TGAGCTATAC ATAGATAATG TGAGGACGGT	1800

TEATTTATCA AATCAGGATG TTATAAACGT TGTAGAGATG GAAATATTTA AAACAGCTTT	1860
GTGGGTTGAA GATAAGTGCT ACATTAGAGA AGATGGTCTT TTTCAGGGCT CTAGTTTATC	1920
TGCTCCGATC GTTGATTTGG TGTATGACGA TCTTCTGGAG TTTTATAGCG AGTTTAAAGC	1980
CAGTCCTAGC CAGGACACAT TAATTTTAAA ACTGGCTGAC GATTTTCCTTA TAATATCAAC	2040
AGACCAACAG CAAGTGATCA ATATCAAAAA GCTTGCCATG GCGGATTTC AAAAATATAA	2100
TGCGAAAGCC AATAGAGACA AAATTTTAGC CGTAAGCTCC CAATCAGATG ATGATACGGT	2160
TATTCAATTT TGTGCAATGC ACATATTTGT TAAAGAATTG GAAGTTTGGA AACATTCAAG	2220
CACAATGAAT AATTTCCATA TCCGTTCGAA ATCTAGTAAA GGGATATTTT GAAGTTTAAT	2280
AGCGCTGTTT AACACTAGAA TCTCTTATAA AACAATTGAC ACAAATTTAA ATTCAACAAA	2340
CACCGTTCTC ATGCAAATTG ATCATGTTGT AAAGAACATT TCGGAATGTT ATAAATCTGC	2400
TTTTAAGGAT CTATCAATTA ATGTTACGCA AAATATGCAA TTTCATTCGT TCTTACAACG	2460
CATCATTGAA ATGACAGTCA GCGGTTGTCC AATTACGAAA TGTGATCCTT TAATCGAGTA	2520
TGAGGTACGA TTCACCATAT TGAATGGATT TTTGGAAAGC CTATCTTCAA ACACATCAAA	2580
ATTTAAAGAT AATATCATTC TTTTGAGAAA GGAAATTCAA CACTTGCAAG C	2631

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 129 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..129
- (D) OTHER INFORMATION: /note= "TRT motifs from human"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Ala	Lys	Phe	Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu
1				5					10					15	
Arg	Ser	Phe	Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu
			20					25					30		
Phe	Phe	Tyr	Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile
			35				40					45			
Arg	Gln	His	Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu
			50			55					60				
Val	Arg	Gln	His	Arg	Glu	Ala	Arg	Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu
65					70					75				80	

Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp  
85 90 95

Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg  
100 105 110

Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg  
115 120 125

Ala

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5544 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(959..1216, 1273..1353, 1425..1543, 1595..1857, 1894..2286, 2326..2396, 2436..2705, 2746..2862, 2914..3083, 3125..3309, 3356..3504, 3546..3759, 3797..4046, 4086..4252, 4296..4392, 4435..4597)
- (D) OTHER INFORMATION: /note= "Schizosaccharomyces pombe telomerase catalytic subunit (TRT)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GGTACCGATT TACTTTCCTT TCTTCATAAG CTAATTGCTT CCTCGAACGC TCCTAAATCT	60
CTGGAAATAT TTTTACAAGA ACTCAATAAC AATACCAAGT CAAATTCCAA TATGAAGGTG	120
TTATTAGTGA TCGATAATAT TTCTATTTTA TCGGTCGTTA CCAAGTATAA GGACAAAAAG	180
AACAACTTCC TTCCCCCTAA AGACTTTTAC TTTATTAATT TACTTTTCAA ATATATTTTCG	240
GGTTCGCTTA CTTTTAATCG TGGTACTGTT TTAGCTGCTA CTTCTAGCCA ACCGCGTGTT	300
TCTACCCCGT CATTGGATAT AGCTCTTGGA GTAGCTCACA GAAATCCTTA CAAATCTTCT	360
GATGAGACTA TATTAGATTC ATTACAGTCC GTGCATATTC TTAACATGGA GCCTTACACT	420
TTAGATGAGT CACGTCGCAT GATGGAGTAT TTGGTATCAT CCAACGTTTG CTTTGAAAAG	480
GTTGATAATT ATTTGCAAAA TCATGTCCTT AGTGGTGGTA ATCCGCGAAA GTTTTTTGTAT	540
GCTTGCACAC GTCTAGCATG ATTGAGATAT TCAAAAATTT CTATCCACTA CAACTCCTTT	600
AACGCGGTTT TATTTTTCTA TTTTCTATTC TCATGTTGTT CCAAATATGT ATCATCTCGT	660
ATTAGGCTTT TTTCCGTTTT ACTCCTGGAA TCGTACCTTT TTTACTATTC CCCCTAATGA	720
ATAATCTAAA TTAGTTTCGC TTATAATTGA TAGTAGTAGA AAGATTGGTG ATTCTACTCG	780

TGTAATGTTA TTAGTTTAAA GATACTTTGC AAAACATTTA TTAGCTATCA TTATATAAAA	840
AAAATCCTAT AATTATAAAT ATTAATCAAT ATTTGCGGTC ACTATTTATT TAAAACGTTA	900
TGATCAGTAG GACACTTTGC ATATATATAG TTATGCTTAA TGGTTACTTG TAACTTGC	958
ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG Met Thr Glu His His Thr Pro Lys Ser Arg Ile Leu Arg Phe Leu Glu 1 5 10 15	1006
AAT CAA TAT GTA TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT Asn Gln Tyr Val Tyr Leu Cys Thr Leu Asn Asp Tyr Val Gln Leu Val 20 25 30	1054
TTG AGA GGG TCG CCG GCA AGC TCG TAT AGC AAT ATA TGC GAA CGC TTG Leu Arg Gly Ser Pro Ala Ser Ser Tyr Ser Asn Ile Cys Glu Arg Leu 35 40 45	1102
AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT CAT TCG ACT GTA Arg Ser Asp Val Gln Thr Ser Phe Ser Ile Phe Leu His Ser Thr Val 50 55 60	1150
GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA Val Gly Phe Asp Ser Lys Pro Asp Glu Gly Val Gln Phe Ser Ser Pro 65 70 75 80	1198
AAA TGC TCA CAG TCA GAG GTATATATAT TTTTGTTTTG ATTTTTTTCT Lys Cys Ser Gln Ser Glu 85	1246
ATTCGGGATA GCTAATATAT GGGCAG CTA ATA GCG AAT GTT GTA AAA CAG ATG Leu Ile Ala Asn Val Val Lys Gln Met 90 95	1299
TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA CTG ATG AAA GGG TTT Phe Asp Glu Ser Phe Glu Arg Arg Arg Asn Leu Leu Met Lys Gly Phe 100 105 110	1347
TCC ATG GTAAGGTATT CTAATTGTGA AATATTTACC TGCAATTACT GTTTCAAAGA Ser Met	1403
GATTGTATTT AACCGATAAA G AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA Asn His Glu Asp Phe Arg Ala Met His Val 115 120	1454
AAC GGA GTA CAA AAT GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA Asn Gly Val Gln Asn Asp Leu Val Ser Thr Phe Pro Asn Tyr Leu Ile 125 130 135	1502
TCT ATA CTT GAG TCA AAA AAT TGG CAA CTT TTG TTA GAA AT Ser Ile Leu Glu Ser Lys Asn Trp Gln Leu Leu Leu Glu Ile 140 145 150	1543
GTAAATACCG GTTAAGATGT TGCGCACTTT GAACAAGACT GACAAGTATA G T ATC Ile	1598
GGC AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG Gly Ser Asp Ala Met His Tyr Leu Leu Ser Lys Gly Ser Ile Phe Glu 155 160 165 170	1646

GCT CTT CCA AAT GAC AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT Ala Leu Pro Asn Asp Asn Tyr Leu Gln Ile Ser Gly Ile Pro Leu Phe 175 180 185	1694
AAA AAT AAT GTG TTT GAG GAA ACT GTG TCA AAA AAA AGA AAG CGA ACC Lys Asn Asn Val Phe Glu Glu Thr Val Ser Lys Lys Arg Lys Arg Thr 190 195 200	1742
ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA GAA GTT TCC Ile Glu Thr Ser Ile Thr Gln Asn Lys Ser Ala Arg Lys Glu Val Ser 205 210 215	1790
TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC Trp Asn Ser Ile Ser Ile Ser Arg Phe Ser Ile Phe Tyr Arg Ser Ser 220 225 230	1838
TAT AAG AAG TTT AAG CAA G GTA ACTAATA CTGTTATCCT TCATAACTAA Tyr Lys Lys Phe Lys Gln 235 240	1887
TTTTAG AT CTA TAT TTT AAC TTA CAC TCT ATT TGT GAT CGG AAC ACA Asp Leu Tyr Phe Asn Leu His Ser Ile Cys Asp Arg Asn Thr 245 250	1934
GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG CAA TTT GGA CTT ATA Val His Met Trp Leu Gln Trp Ile Phe Pro Arg Gln Phe Gly Leu Ile 255 260 265 270	1982
AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA TCA Asn Ala Phe Gln Val Lys Gln Leu His Lys Val Ile Pro Leu Val Ser 275 280 285	2030
CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT Gln Ser Thr Val Val Pro Lys Arg Leu Leu Lys Val Tyr Pro Leu Ile 290 295 300	2078
GAA CAA ACA GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC Glu Gln Thr Ala Lys Arg Leu His Arg Ile Ser Leu Ser Lys Val Tyr 305 310 315	2126
AAC CAT TAT TGC CCA TAT ATT GAC ACC CAC GAT GAT GAA AAA ATC CTT Asn His Tyr Cys Pro Tyr Ile Asp Thr His Asp Asp Glu Lys Ile Leu 320 325 330	2174
AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG TTT CTT CGA TCC ATT Ser Tyr Ser Leu Lys Pro Asn Gln Val Phe Ala Phe Leu Arg Ser Ile 335 340 345 350	2222
CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA TTT Leu Val Arg Val Phe Pro Lys Leu Ile Trp Gly Asn Gln Arg Ile Phe 355 360 365	2270
GAG ATA ATA TTA AAA G GTATTGTATA AAATTTATTA CCACTAACGA TTTTACCAG AC Glu Ile Ile Leu Lys 370 Asp	2327
CTC GAA ACT TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT Leu Glu Thr Phe Leu Lys Leu Ser Arg Tyr Glu Ser Phe Ser Leu His 375 380 385	2375
TAT TTA ATG AGT AAC ATA AAG GTAATATGCC AAATTTTTTTT ACCATTAATT Tyr Leu Met Ser Asn Ile Lys	2426

390	395	
AACAATCAG ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA AAA AGG TCA		2474
Ile Ser Glu Ile Glu Trp Leu Val Leu Gly Lys Arg Ser		
400	405	
AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT		2522
Asn Ala Lys Met Cys Leu Ser Asp Phe Glu Lys Arg Lys Gln Ile Phe		
410	415	420
GCG GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA		2570
Ala Glu Phe Ile Tyr Trp Leu Tyr Asn Ser Phe Ile Ile Pro Ile Leu		
425	430	435 440
CAA TCT TTT TTT TAT ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT		2618
Gln Ser Phe Phe Tyr Ile Thr Glu Ser Ser Asp Leu Arg Asn Arg Thr		
445	450	455
GTT TAT TTT AGA AAA GAT ATT TGG AAA CTC TTG TGC CGA CCC TTT ATT		2666
Val Tyr Phe Arg Lys Asp Ile Trp Lys Leu Leu Cys Arg Pro Phe Ile		
460	465	470
ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG GTATTTTAAA		2715
Thr Ser Met Lys Met Glu Ala Phe Glu Lys Ile Asn Glu		
475	480	485
GTATTTTTTG CAAAAAGCTA ATATTTTCAG AAC AAT GTT AGG ATG GAT ACT CAG		2769
Asn Asn Val Arg Met Asp Thr Gln		
490		
AAA ACT ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT		2817
Lys Thr Thr Leu Pro Pro Ala Val Ile Arg Leu Leu Pro Lys Lys Asn		
495	500	505
ACC TTT CGT CTC ATT ACG AAT TTA AGA AAA AGA TTC TTA ATA AAG		2862
Thr Phe Arg Leu Ile Thr Asn Leu Arg Lys Arg Phe Leu Ile Lys		
510	515	520
GTATTAATTT TTGGTCATCA ATGTACTTTA CTTCTAATCT ATTATTAGCA G ATG GGT		2919
Met Gly		
525		
TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG		2967
Ser Asn Lys Lys Met Leu Val Ser Thr Asn Gln Thr Leu Arg Pro Val		
530	535	540
GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA		3015
Ala Ser Ile Leu Lys His Leu Ile Asn Glu Glu Ser Ser Gly Ile Pro		
545	550	555
TTT AAC TTG GAG GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT		3063
Phe Asn Leu Glu Val Tyr Met Lys Leu Leu Thr Phe Lys Lys Asp Leu		
560	565	570
CTT AAG CAC CGA ATG TTT GG GTAATTATAT AATGCGCGAT TCCTCATTAT		3113
Leu Lys His Arg Met Phe Gly		
575	580	
TAATTTTGCA G G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA AAA TCC		3161
Arg Lys Lys Tyr Phe Val Arg Ile Asp Ile Lys Ser		
585	590	

TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG Cys Tyr Asp Arg Ile Lys Gln Asp Leu Met Phe Arg Ile Val Lys Lys 595 600 605	3209
AAA CTC AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT Lys Leu Lys Asp Pro Glu Phe Val Ile Arg Lys Tyr Ala Thr Ile His 610 615 620 625	3257
GCA ACA AGT GAC CGA GCT ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC Ala Thr Ser Asp Arg Ala Thr Lys Asn Phe Val Ser Glu Ala Phe Ser 630 635 640	3305
TAT T GTAAGTTTAT TTTTTCATTG GAATTTTTTA ACAAATTCTT TTTTAG TT Tyr Phe	3357
GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA Asp Met Val Pro Phe Glu Lys Val Val Gln Leu Leu Ser Met Lys Thr 645 650 655	3405
TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT Ser Asp Thr Leu Phe Val Asp Phe Val Asp Tyr Trp Thr Lys Ser Ser 660 665 670 675	3453
TCT GAA ATT TTT AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT Ser Glu Ile Phe Lys Met Leu Lys Glu His Leu Ser Gly His Ile Val 680 685 690	3501
AAG GTATACCAAT TGTTGAATTG TAATAACACT AATGAAACTA G ATA GGA AAT Lys Ile Gly Asn 695	3554
TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA ATT CTG TCA Ser Gln Tyr Leu Gln Lys Val Gly Ile Pro Gln Gly Ser Ile Leu Ser 700 705 710	3602
TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA Ser Phe Leu Cys His Phe Tyr Met Glu Asp Leu Ile Asp Glu Tyr Leu 715 720 725	3650
TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT Ser Phe Thr Lys Lys Lys Gly Ser Val Leu Leu Arg Val Val Asp Asp 730 735 740	3698
TTC CTC TTT ATA ACA GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT Phe Leu Phe Ile Thr Val Asn Lys Lys Asp Ala Lys Lys Phe Leu Asn 745 750 755	3746
TTA TCT TTA AGA G GTGAGTTGCT GTCATTCCTA AGTTCTAACC GTTGAAG GA Leu Ser Leu Arg Gly	3798
TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA ATA AAC Phe Glu Lys His Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile Asn 765 770 775 780	3846
TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC Phe Glu Asn Ser Asn Gly Ile Ile Asn Asn Thr Phe Phe Asn Glu Ser 785 790 795	3894
AAG AAA AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT Lys Lys Arg Met Pro Phe Phe Gly Phe Ser Val Asn Met Arg Ser Leu	3942



800	805	810	
GAT ACA TTG TTA GCA TGT CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT Asp Thr Leu Leu Ala Cys Pro Lys Ile Asp Glu Ala Leu Phe Asn Ser 815 820 825			3990
ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG AAA TCT TTT TTT TAC AAA Thr Ser Val Glu Leu Thr Lys His Met Gly Lys Ser Phe Phe Tyr Lys 830 835 840			4038
ATT CTA AG GTATACTGTG TAACTGAATA ATAGCTGACA AATAATCAG A TCG Ile Leu Arg Ser 845			4089
AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA Ser Leu Ala Ser Phe Ala Gln Val Phe Ile Asp Ile Thr His Asn Ser 850 855 860			4137
AAA TTC AAT TCT TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT Lys Phe Asn Ser Cys Cys Asn Ile Tyr Arg Leu Gly Tyr Ser Met Cys 865 870 875 880			4185
ATG AGA GCA CAA GCA TAC TTA AAA AGG ATG AAG GAT ATA TTT ATT CCC Met Arg Ala Gln Ala Tyr Leu Lys Arg Met Lys Asp Ile Phe Ile Pro 885 890 895			4233
CAA AGA ATG TTC ATA ACG G GTGAGTACTT ATTTTAACTA GAAAAGTCAT Gln Arg Met Phe Ile Thr 900			4282
TAATTAACCT TAG AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA Asp Leu Leu Asn Val Ile Gly Arg Lys Ile Trp Lys 905 910			4330
AAG TTG GCC GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT Lys Leu Ala Glu Ile Leu Gly Tyr Thr Ser Arg Arg Phe Leu Ser Ser 915 920 925 930			4378
GCA GAA GTC AAA TG GTACGTGTCG GTCTCGAGAC TTCAGCAATA TTGACACATC Ala Glu Val Lys Trp 935			4432
AG G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA CCC TCT TTC AAA Leu Phe Cys Leu Gly Met Arg Asp Gly Leu Lys Pro Ser Phe Lys 940 945 950			4480
TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT Tyr His Pro Cys Phe Glu Gln Leu Ile Tyr Gln Phe Gln Ser Leu Thr 955 960 965			4528
GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA Asp Leu Ile Lys Pro Leu Arg Pro Val Leu Arg Gln Val Leu Phe Leu 970 975 980			4576
CAT AGA AGA ATA GCT GAT TAATGTCATT TTCAATTTAT TATATACATC His Arg Arg Ile Ala Asp 985			4624
CTTTATTACT GGTGTCTTAA ACAATATTAT TACTAAGTAT AGCTGACCCC CAAAGCAAGC			4684
ATACTATAGG ATTTCTAGTA AAGTAAAATT AATCTCGTTA TTAGTTTTGA TTGACTTGTC			4744

TTTATCCTTA TACTTTTAAG AAAGATTGAC AGTGGTTGCT GACTACTGCC CACATGCCCA	4804
TTAAACGGGA GTGGTTAAAC ATTAAAAGTA ATACATGAGG CTAATCTCCT TTCATTTAGA	4864
ATAAGGAAAG TGGTTTTCTA TAATGAATAA TGCCCGCACT AATGCAAAAA GACGAAGATT	4924
ATCTTCTAAA CAAGGGGGAT TAAGCATATC CGAAGGAAAA GAGAGTAATA TACCCAGTGT	4984
TGTTGAAGAA AGCAAGGATA ATTTGGAACA AGCTTCTGCA GATGACAGGC TAAATTTTGG	5044
TGACCGAATT TTGGTAAAAG CCCAGGTTA TCCATGGTGG CCGGCCTTGC TACTGAGACG	5104
AAAAGAAACT AAGGATAGTT TGAATACTAA TAGCTCATTT AATGTCTTAT ATAAGGTTTT	5164
GTTTTTTCCT GACTTCAATT TTGCATGGGT GAAAAGAAAT AGTGTTAAGC CATTATTGGA	5224
TTCCGAAATA GCCAAATTTC TTGGTTCCTC AAAGCGGAAG TCTAAAGAAC TTATTGAAGC	5284
TTATGAGGCT TCAAAAACTC CTCCTGATTT AAAGGAGGAA TCTTCCACCG ATGAGGAAAT	5344
GGATAGCTTA TCAGCTGCTG AGGAGAAGCC TAATTTTTTG CAAAAAGAA AATATCATTG	5404
GGAGACATCT CTTGATGAAT CAGATGCGGA GAGTATCTCC AGCGGATCCT TGATGTCAAT	5464
AACTTCTATT TCTGAAATGT ATGGTCCTAC TGTCGCTTCG ACTTCTCGTA GCTCTACGCA	5524
GTTAAGTGAC CAAAGGTACC	5544

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 988 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Met	Thr	Glu	His	His	Thr	Pro	Lys	Ser	Arg	Ile	Leu	Arg	Phe	Leu	Glu	1	5	10	15
Asn	Gln	Tyr	Val	Tyr	Leu	Cys	Thr	Leu	Asn	Asp	Tyr	Val	Gln	Leu	Val	20	25	30	
Leu	Arg	Gly	Ser	Pro	Ala	Ser	Ser	Tyr	Ser	Asn	Ile	Cys	Glu	Arg	Leu	35	40	45	
Arg	Ser	Asp	Val	Gln	Thr	Ser	Phe	Ser	Ile	Phe	Leu	His	Ser	Thr	Val	50	55	60	
Val	Gly	Phe	Asp	Ser	Lys	Pro	Asp	Glu	Gly	Val	Gln	Phe	Ser	Ser	Pro	65	70	75	80
Lys	Cys	Ser	Gln	Ser	Glu	Leu	Ile	Ala	Asn	Val	Val	Lys	Gln	Met	Phe	85	90	95	
Asp	Glu	Ser	Phe	Glu	Arg	Arg	Arg	Asn	Leu	Leu	Met	Lys	Gly	Phe	Ser	100	105	110	
Met	Asn	His	Glu	Asp	Phe	Arg	Ala	Met	His	Val	Asn	Gly	Val	Gln	Asn				

115					120					125					
Asp	Leu	Val	Ser	Thr	Phe	Pro	Asn	Tyr	Leu	Ile	Ser	Ile	Leu	Glu	Ser
130						135					140				
Lys	Asn	Trp	Gln	Leu	Leu	Leu	Glu	Ile	Ile	Gly	Ser	Asp	Ala	Met	His
145					150					155					160
Tyr	Leu	Leu	Ser	Lys	Gly	Ser	Ile	Phe	Glu	Ala	Leu	Pro	Asn	Asp	Asn
				165					170					175	
Tyr	Leu	Gln	Ile	Ser	Gly	Ile	Pro	Leu	Phe	Lys	Asn	Asn	Val	Phe	Glu
			180					185					190		
Glu	Thr	Val	Ser	Lys	Lys	Arg	Lys	Arg	Thr	Ile	Glu	Thr	Ser	Ile	Thr
		195					200					205			
Gln	Asn	Lys	Ser	Ala	Arg	Lys	Glu	Val	Ser	Trp	Asn	Ser	Ile	Ser	Ile
	210					215					220				
Ser	Arg	Phe	Ser	Ile	Phe	Tyr	Arg	Ser	Ser	Tyr	Lys	Lys	Phe	Lys	Gln
225					230					235					240
Asp	Leu	Tyr	Phe	Asn	Leu	His	Ser	Ile	Cys	Asp	Arg	Asn	Thr	Val	His
				245					250					255	
Met	Trp	Leu	Gln	Trp	Ile	Phe	Pro	Arg	Gln	Phe	Gly	Leu	Ile	Asn	Ala
			260					265					270		
Phe	Gln	Val	Lys	Gln	Leu	His	Lys	Val	Ile	Pro	Leu	Val	Ser	Gln	Ser
		275					280					285			
Thr	Val	Val	Pro	Lys	Arg	Leu	Leu	Lys	Val	Tyr	Pro	Leu	Ile	Glu	Gln
	290					295					300				
Thr	Ala	Lys	Arg	Leu	His	Arg	Ile	Ser	Leu	Ser	Lys	Val	Tyr	Asn	His
305					310					315					320
Tyr	Cys	Pro	Tyr	Ile	Asp	Thr	His	Asp	Asp	Glu	Lys	Ile	Leu	Ser	Tyr
				325					330					335	
Ser	Leu	Lys	Pro	Asn	Gln	Val	Phe	Ala	Phe	Leu	Arg	Ser	Ile	Leu	Val
			340					345					350		
Arg	Val	Phe	Pro	Lys	Leu	Ile	Trp	Gly	Asn	Gln	Arg	Ile	Phe	Glu	Ile
		355					360					365			
Ile	Leu	Lys	Asp	Leu	Glu	Thr	Phe	Leu	Lys	Leu	Ser	Arg	Tyr	Glu	Ser
	370					375					380				
Phe	Ser	Leu	His	Tyr	Leu	Met	Ser	Asn	Ile	Lys	Ile	Ser	Glu	Ile	Glu
385					390					395					400
Trp	Leu	Val	Leu	Gly	Lys	Arg	Ser	Asn	Ala	Lys	Met	Cys	Leu	Ser	Asp
				405					410				415		
Phe	Glu	Lys	Arg	Lys	Gln	Ile	Phe	Ala	Glu	Phe	Ile	Tyr	Trp	Leu	Tyr
			420					425					430		
Asn	Ser	Phe	Ile	Ile	Pro	Ile	Leu	Gln	Ser	Phe	Phe	Tyr	Ile	Thr	Glu
		435					440					445			

Ser Ser Asp Leu Arg Asn Arg Thr Val Tyr Phe Arg Lys Asp Ile Trp  
 450 455 460  
 Lys Leu Leu Cys Arg Pro Phe Ile Thr Ser Met Lys Met Glu Ala Phe  
 465 470 475 480  
 Glu Lys Ile Asn Glu Asn Asn Val Arg Met Asp Thr Gln Lys Thr Thr  
 485 490 495  
 Leu Pro Pro Ala Val Ile Arg Leu Leu Pro Lys Lys Asn Thr Phe Arg  
 500 505 510  
 Leu Ile Thr Asn Leu Arg Lys Arg Phe Leu Ile Lys Met Gly Ser Asn  
 515 520 525  
 Lys Lys Met Leu Val Ser Thr Asn Gln Thr Leu Arg Pro Val Ala Ser  
 530 535 540  
 Ile Leu Lys His Leu Ile Asn Glu Glu Ser Ser Gly Ile Pro Phe Asn  
 545 550 555 560  
 Leu Glu Val Tyr Met Lys Leu Leu Thr Phe Lys Lys Asp Leu Leu Lys  
 565 570 575  
 His Arg Met Phe Gly Arg Lys Lys Tyr Phe Val Arg Ile Asp Ile Lys  
 580 585 590  
 Ser Cys Tyr Asp Arg Ile Lys Gln Asp Leu Met Phe Arg Ile Val Lys  
 595 600 605  
 Lys Lys Leu Lys Asp Pro Glu Phe Val Ile Arg Lys Tyr Ala Thr Ile  
 610 615 620  
 His Ala Thr Ser Asp Arg Ala Thr Lys Asn Phe Val Ser Glu Ala Phe  
 625 630 635 640  
 Ser Tyr Phe Asp Met Val Pro Phe Glu Lys Val Val Gln Leu Leu Ser  
 645 650 655  
 Met Lys Thr Ser Asp Thr Leu Phe Val Asp Phe Val Asp Tyr Trp Thr  
 660 665 670  
 Lys Ser Ser Ser Glu Ile Phe Lys Met Leu Lys Glu His Leu Ser Gly  
 675 680 685  
 His Ile Val Lys Ile Gly Asn Ser Gln Tyr Leu Gln Lys Val Gly Ile  
 690 695 700  
 Pro Gln Gly Ser Ile Leu Ser Ser Phe Leu Cys His Phe Tyr Met Glu  
 705 710 715 720  
 Asp Leu Ile Asp Glu Tyr Leu Ser Phe Thr Lys Lys Lys Gly Ser Val  
 725 730 735  
 Leu Leu Arg Val Val Asp Asp Phe Leu Phe Ile Thr Val Asn Lys Lys  
 740 745 750  
 Asp Ala Lys Lys Phe Leu Asn Leu Ser Leu Arg Gly Phe Glu Lys His  
 755 760 765  
 Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile Asn Phe Glu Asn Ser  
 770 775 780

Asn Gly Ile Ile Asn Asn Thr Phe Phe Asn Glu Ser Lys Lys Arg Met  
 785 790 795 800  
 Pro Phe Phe Gly Phe Ser Val Asn Met Arg Ser Leu Asp Thr Leu Leu  
 805 810 815  
 Ala Cys Pro Lys Ile Asp Glu Ala Leu Phe Asn Ser Thr Ser Val Glu  
 820 825 830  
 Leu Thr Lys His Met Gly Lys Ser Phe Phe Tyr Lys Ile Leu Arg Ser  
 835 840 845  
 Ser Leu Ala Ser Phe Ala Gln Val Phe Ile Asp Ile Thr His Asn Ser  
 850 855 860  
 Lys Phe Asn Ser Cys Cys Asn Ile Tyr Arg Leu Gly Tyr Ser Met Cys  
 865 870 875 880  
 Met Arg Ala Gln Ala Tyr Leu Lys Arg Met Lys Asp Ile Phe Ile Pro  
 885 890 895  
 Gln Arg Met Phe Ile Thr Asp Leu Leu Asn Val Ile Gly Arg Lys Ile  
 900 905 910  
 Trp Lys Lys Leu Ala Glu Ile Leu Gly Tyr Thr Ser Arg Arg Phe Leu  
 915 920 925  
 Ser Ser Ala Glu Val Lys Trp Leu Phe Cys Leu Gly Met Arg Asp Gly  
 930 935 940  
 Leu Lys Pro Ser Phe Lys Tyr His Pro Cys Phe Glu Gln Leu Ile Tyr  
 945 950 955 960  
 Gln Phe Gln Ser Leu Thr Asp Leu Ile Lys Pro Leu Arg Pro Val Leu  
 965 970 975  
 Arg Gln Val Leu Phe Leu His Arg Arg Ile Ala Asp  
 980 985

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /mod\_base= OTHER  
/note= "N = guanosine modified by a biotin group"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

NCCTATTTT TYTAYNNNAC NGA

23

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Phe Phe Tyr Xaa Thr Glu  
1 5

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CCAGATATNA DNARRAARTC RTC

23

(2) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Phe, Ile or Leu"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Asp Asp Phe Leu Xaa Ile  
1 5

(2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

ACAATGMGNH TNHTNCCNAA RAA

23

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2..3
- (D) OTHER INFORMATION: /product= "OTHER"  
/note= "Xaa = Leu or Ile"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Arg Xaa Xaa Pro Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

ACGAATCKNG GDATNSWRTC RTARCA

26

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Cys Tyr Asp Ser Ile Pro Arg  
1 5

(2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CAATTCTCRT ARCANSWYTT DATRTC

26

(2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 7 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Asp Ile Lys Ser Cys Tyr Asp  
1                    5

(2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 269 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

GATTACTCCC GAAGAAAGGA TCTTTCCGTC CAATCATGAC TTTCTTAAGA AAGGACAAGC	60
AAAAAAATAT TAAGTTAAAT CTAAATTAAA TTCTAATGGA TAGCCAACCTT GTGTTTAGGA	120
ATTTAAAAGA CATGCTGGGA TAAAAGATAG GATACTCAGT CTTTGATAAT AAACAAATTT	180
CAGAAAAATT TGCCTAATTC ATAGAGAAAT GGAAAAATAA AGGAAGACCT CAGCTATATT	240
ATGTCACTCT AGACATAAAG ACTTGCTAC	269

(2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 474 base pairs  
    (B) TYPE: nucleic acid



(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

AAACACAAGG AAGGAAGTCA AATATTCTAT TACCGTAAAC CAATATGGAA ATTAGTGAGT	60
AAATTAACTA TTGTCAAAGT AAGAATTTAG TTTTCTGAAA AGAATAAATA AATGAAAAAT	120
AATTTTATC AAAAAATTTA GCTTGAAGAG GAGAATTGG AAAAAAGTTGA AGAAAAATTG	180
ATACCAGAAG ATTCATTTTA GAAATACCCCT CAAGGAAAGC TAAGGATTAT ACCTAAAAAA	240
GGATCTTTCC GTCCAATCAT GACTTTCTTA AGAAAGGACA AGCAAAAAAA TATTAAGTTA	300
AATCTAAATT AAATTCTAAT GGATAGCCAA CTTGTGTTTA GGAATTTAAA AGACATGCTG	360
GGATAAAAGA TAGGATACTC AGTCTTTGAT AATAAACAAA TTTCAGAAAA ATTTGCCTAA	420
TTCATAGAGA AATGGAAAAA TAAAGGAAGA CCTCAGCTAT ATTATGTCAC TCTA	474

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 158 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Lys His Lys Glu Gly Ser Gln Ile Phe Tyr Tyr Arg Lys Pro Ile Trp	
1 5 10 15	
Lys Leu Val Ser Lys Leu Thr Ile Val Lys Val Arg Ile Gln Phe Ser	
20 25 30	
Glu Lys Asn Lys Gln Met Lys Asn Asn Phe Tyr Gln Lys Ile Gln Leu	
35 40 45	
Glu Glu Glu Asn Leu Glu Lys Val Glu Glu Lys Leu Ile Pro Glu Asp	
50 55 60	
Ser Phe Gln Lys Tyr Pro Gln Gly Lys Leu Arg Ile Ile Pro Lys Lys	
65 70 75 80	
Gly Ser Phe Arg Pro Ile Met Thr Phe Leu Arg Lys Asp Lys Gln Lys	
85 90 95	
Asn Ile Lys Leu Asn Leu Asn Gln Ile Leu Met Asp Ser Gln Leu Val	
100 105 110	
Phe Arg Asn Leu Lys Asp Met Leu Gly Gln Lys Ile Gly Tyr Ser Val	
115 120 125	
Phe Asp Asn Lys Gln Ile Ser Glu Lys Phe Ala Gln Phe Ile Glu Lys	

130	135	140
Trp Lys Asn Lys Gly Arg Pro Gln Leu Tyr Tyr Val Thr Leu		
145	150	155

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 157 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Phe	Phe	Tyr	Cys	Thr	Glu	Ile	Ser	Ser	Thr	Val	Thr	Ile	Val	Tyr	Phe
1				5					10					15	
Arg	His	Asp	Thr	Trp	Asn	Lys	Leu	Ile	Thr	Pro	Phe	Ile	Val	Glu	Tyr
			20					25					30		
Phe	Lys	Thr	Tyr	Leu	Val	Glu	Asn	Asn	Val	Cys	Arg	Asn	His	Asn	Ser
			35				40					45			
Tyr	Thr	Leu	Ser	Asn	Phe	Asn	His	Ser	Lys	Met	Arg	Ile	Ile	Pro	Lys
			50			55					60				
Lys	Ser	Asn	Asn	Glu	Phe	Arg	Ile	Ile	Ala	Ile	Pro	Cys	Arg	Gly	Ala
65					70				75					80	
Asp	Glu	Glu	Glu	Phe	Thr	Ile	Tyr	Lys	Glu	Asn	His	Lys	Asn	Ala	Ile
				85					90					95	
Gln	Pro	Thr	Gln	Lys	Ile	Leu	Glu	Tyr	Leu	Arg	Asn	Lys	Arg	Pro	Thr
			100					105					110		
Ser	Phe	Thr	Lys	Ile	Tyr	Ser	Pro	Thr	Gln	Ile	Ala	Asp	Arg	Ile	Lys
			115				120					125			
Glu	Phe	Lys	Gln	Arg	Leu	Leu	Lys	Lys	Phe	Asn	Asn	Val	Leu	Pro	Glu
	130					135					140				
Leu	Tyr	Phe	Met	Lys	Phe	Asp	Val	Lys	Ser	Cys	Tyr	Asp			
145					150					155					

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 155 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Phe	Phe	Tyr	Val	Thr	Glu	Gln	Gln	Lys	Ser	Tyr	Ser	Lys	Thr	Tyr	Tyr	1	5	10	15
Tyr	Arg	Lys	Asn	Ile	Trp	Asp	Val	Ile	Met	Lys	Met	Ser	Ile	Ala	Asp	20	25	30	
Leu	Lys	Lys	Glu	Thr	Leu	Ala	Glu	Val	Gln	Glu	Lys	Glu	Val	Glu	Glu	35	40	45	
Trp	Lys	Lys	Ser	Leu	Gly	Phe	Ala	Pro	Gly	Lys	Leu	Arg	Leu	Ile	Pro	50	55	60	
Lys	Lys	Thr	Thr	Phe	Arg	Pro	Ile	Met	Thr	Phe	Asn	Lys	Lys	Ile	Val	65	70	75	80
Asn	Ser	Asp	Arg	Lys	Thr	Thr	Lys	Leu	Thr	Thr	Asn	Thr	Lys	Leu	Leu	85	90	95	
Asn	Ser	His	Leu	Met	Leu	Lys	Thr	Leu	Lys	Asn	Arg	Met	Phe	Lys	Asp	100	105	110	
Pro	Phe	Gly	Phe	Ala	Val	Phe	Asn	Tyr	Asp	Asp	Val	Met	Lys	Lys	Tyr	115	120	125	
Glu	Glu	Phe	Val	Cys	Lys	Trp	Lys	Gln	Val	Gly	Gln	Pro	Lys	Leu	Phe	130	135	140	
Phe	Ala	Thr	Met	Asp	Ile	Glu	Lys	Cys	Tyr	Asp	145	150	155						

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 158 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Lys	His	Lys	Glu	Gly	Ser	Gln	Ile	Phe	Tyr	Tyr	Arg	Lys	Pro	Ile	Trp	1	5	10	15
Lys	Leu	Val	Ser	Lys	Leu	Thr	Ile	Val	Lys	Val	Arg	Ile	Gln	Phe	Ser	20	25	30	
Glu	Lys	Asn	Lys	Gln	Met	Lys	Asn	Asn	Phe	Tyr	Gln	Lys	Ile	Gln	Leu	35	40	45	
Glu	Glu	Glu	Asn	Leu	Glu	Lys	Val	Glu	Glu	Lys	Leu	Ile	Pro	Glu	Asp	50	55	60	
Ser	Phe	Gln	Lys	Tyr	Pro	Gln	Gly	Lys	Leu	Arg	Ile	Ile	Pro	Lys	Lys	65	70	75	80
Gly	Ser	Phe	Arg	Pro	Ile	Met	Thr	Phe	Leu	Arg	Lys	Asp	Lys	Gln	Lys	85	90	95	

Asn Ile Lys Leu Asn Leu Asn Gln Ile Leu Met Asp Ser Gln Leu Val  
100 105 110  
Phe Arg Asn Leu Lys Asp Met Leu Gly Gln Lys Ile Gly Tyr Ser Val  
115 120 125  
Phe Asp Asn Lys Gln Ile Ser Glu Lys Phe Ala Gln Phe Ile Glu Lys  
130 135 140  
Trp Lys Asn Lys Gly Arg Pro Gln Leu Tyr Tyr Val Thr Leu  
145 150 155

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1007 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Glu Val Asp Val Asp Asn Gln Ala Asp Asn His Gly Ile His Ser Ala  
1 5 10 15  
Leu Lys Thr Cys Glu Glu Ile Lys Glu Ala Lys Thr Leu Tyr Ser Trp  
20 25 30  
Ile Gln Lys Val Ile Arg Cys Arg Asn Gln Ser Gln Ser His Tyr Lys  
35 40 45  
Asp Leu Glu Asp Ile Lys Ile Phe Ala Gln Thr Asn Ile Val Ala Thr  
50 55 60  
Pro Arg Asp Tyr Asn Glu Glu Asp Phe Lys Val Ile Ala Arg Lys Glu  
65 70 75 80  
Val Phe Ser Thr Gly Leu Met Ile Glu Leu Ile Asp Lys Cys Leu Val  
85 90 95  
Glu Leu Leu Ser Ser Ser Asp Val Ser Asp Arg Gln Lys Leu Gln Cys  
100 105 110  
Phe Gly Phe Gln Leu Lys Gly Asn Gln Leu Ala Lys Thr His Leu Leu  
115 120 125  
Thr Ala Leu Ser Thr Gln Lys Gln Tyr Phe Phe Gln Asp Glu Trp Asn  
130 135 140  
Gln Val Arg Ala Met Ile Gly Asn Glu Leu Phe Arg His Leu Tyr Thr  
145 150 155 160  
Lys Tyr Leu Ile Phe Gln Arg Thr Ser Glu Gly Thr Leu Val Gln Phe  
165 170 175  
Cys Gly Asn Asn Val Phe Asp His Leu Lys Val Asn Asp Lys Phe Asp  
180 185 190



Asn Lys Lys Ile Val Asn Ser Asp Arg Lys Thr Thr Lys Leu Thr Thr  
 530 535 540  
 Asn Thr Lys Leu Leu Asn Ser His Leu Met Leu Lys Thr Leu Lys Asn  
 545 550 555 560  
 Arg Met Phe Lys Asp Pro Phe Gly Phe Ala Val Phe Asn Tyr Asp Asp  
 565 570 575  
 Val Met Lys Lys Tyr Glu Glu Phe Val Cys Lys Trp Lys Gln Val Gly  
 580 585 590  
 Gln Pro Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr Asp  
 595 600 605  
 Ser Val Asn Arg Glu Lys Leu Ser Thr Phe Leu Lys Thr Thr Lys Leu  
 610 615 620  
 Leu Ser Ser Asp Phe Trp Ile Met Thr Ala Gln Ile Leu Lys Arg Lys  
 625 630 635 640  
 Asn Asn Ile Val Ile Asp Ser Lys Asn Phe Arg Lys Lys Glu Met Lys  
 645 650 655  
 Asp Tyr Phe Arg Gln Lys Phe Gln Lys Ile Ala Leu Glu Gly Gly Gln  
 660 665 670  
 Tyr Pro Thr Leu Phe Ser Val Leu Glu Asn Glu Gln Asn Asp Leu Asn  
 675 680 685  
 Ala Lys Lys Thr Leu Ile Val Glu Ala Lys Gln Arg Asn Tyr Phe Lys  
 690 695 700  
 Lys Asp Asn Leu Leu Gln Pro Val Ile Asn Ile Cys Gln Tyr Asn Tyr  
 705 710 715 720  
 Ile Asn Phe Asn Gly Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro Gln  
 725 730 735  
 Gly Leu Cys Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr Leu  
 740 745 750  
 Glu Glu Ser Ser Leu Gly Phe Leu Arg Asp Glu Ser Met Asn Pro Glu  
 755 760 765  
 Asn Pro Asn Val Asn Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu  
 770 775 780  
 Ile Thr Thr Gln Glu Asn Asn Ala Val Leu Phe Ile Glu Lys Leu Ile  
 785 790 795 800  
 Asn Val Ser Arg Glu Asn Gly Phe Lys Phe Asn Met Lys Lys Leu Gln  
 805 810 815  
 Thr Ser Phe Pro Leu Ser Pro Ser Lys Phe Ala Lys Tyr Gly Met Asp  
 820 825 830  
 Ser Val Glu Glu Gln Asn Ile Val Gln Asp Tyr Cys Asp Trp Ile Gly  
 835 840 845  
 Ile Ser Ile Asp Met Lys Thr Leu Ala Leu Met Pro Asn Ile Asn Leu  
 850 855 860

Arg Ile Glu Gly Ile Leu Cys Thr Leu Asn Leu Asn Met Gln Thr Lys  
 865 870 875 880  
 Lys Ala Ser Met Trp Leu Lys Lys Lys Leu Lys Ser Phe Leu Met Asn  
 885 890 895  
 Asn Ile Thr His Tyr Phe Arg Lys Thr Ile Thr Thr Glu Asp Phe Ala  
 900 905 910  
 Asn Lys Thr Leu Asn Lys Leu Phe Ile Ser Gly Gly Tyr Lys Tyr Met  
 915 920 925  
 Gln Cys Ala Lys Glu Tyr Lys Asp His Phe Lys Lys Asn Leu Ala Met  
 930 935 940  
 Ser Ser Met Ile Asp Leu Glu Val Ser Lys Ile Ile Tyr Ser Val Thr  
 945 950 955 960  
 Arg Ala Phe Phe Lys Tyr Leu Val Cys Asn Ile Lys Asp Thr Ile Phe  
 965 970 975  
 Gly Glu Glu His Tyr Pro Asp Phe Phe Leu Ser Thr Leu Lys His Phe  
 980 985 990  
 Ile Glu Ile Phe Ser Thr Lys Lys Tyr Ile Phe Asn Arg Val Cys  
 995 1000 1005

(2) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GTGAAGGCAC TGTT CAGCG

19

(2) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

GTGGATGATT TCTTGT TGG

19

(2) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

ATGCTCCTGC GTTTGGTGG

19

(2) INFORMATION FOR SEQ ID NO:90:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CTGGACACTC AGCCCTTGG

19

(2) INFORMATION FOR SEQ ID NO:91:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

GGCAGGTGTG CTGGACACT

19

(2) INFORMATION FOR SEQ ID NO:92:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

TTTGATGATG CTGGCGATG

19

(2) INFORMATION FOR SEQ ID NO:93:

- (i) SEQUENCE CHARACTERISTICS:



- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GGGGCTCGTC TTCTACAGG

19

(2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CAGCAGGAGG ATCTTGTAG

19

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

TGACCCCAGG AGTGGCAGG

19

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

TCAAGCTGAC TCGACACCG

19

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CGGCGTGACA GGGCTGC

17

(2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GCTGAAGGCT GAGTGTCC

18

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

TAGTCCATGT TCACAATCG

19

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2171 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 22..1716
- (D) OTHER INFORMATION: /note= "EcoRI-NotI insert of clone 712562 encoding 63 kDa

telomerase protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GCCAAGTTCC TGC	ACTGGCT	GATGAGTGTG	TACGTCGTCG	AGCTGCTCAG	GTCTTTCTTT	60
TATGTCACGG AG	ACCACGTT	TCAAAAGAAC	AGGCTCTTTT	TCTACCGGAA	GAGTGTCTGG	120
AGCAAGTTGC AA	AGCATTGG	AATCAGACAG	CACTTGAAGA	GGGTGCAGCT	GCGGGAGCTG	180
TCGGAAGCAG AG	GTCAAGCA	GCATCGGGAA	GCCAGGCCCG	CCCTGCTGAC	GTCCAGACTC	240
CGCTTCATCC CC	AAGCCTGA	CGGGCTGCGG	CCGATTGTGA	ACATGGACTA	CGTCGTGGGA	300
GCCAGAACGT TC	CGCAGAGA	AAAGARGGCC	GAGCGTCTCA	CCTCGAGGGT	GAAGGCACTG	360
TTCAGCGTGC TC	AACTACGA	GCGGGCGCGG	CGCCCCGGCC	TCCTGGGCGC	CTCTGTGCTG	420
GGCCTGGACG AT	ATCCACAG	GGCCTGGCGC	ACCTTCGTGC	TGCGTGTGCG	GGCCCAGGAC	480
CCGCCGCCTG AG	CTGTACTT	TGTCAAGGTG	GATGTGACGG	GCGCGTACGA	CACCATCCCC	540
CAGGACAGGC TC	ACGGAGGT	CATCGCCAGC	ATCATCAAAC	CCCAGAACAC	GTA	600
CGTCGGTATG CC	GTGGTCCA	GAAGGCCGCC	ATGGGCACGT	CCGCAAGGCC	TTCAAGAGCC	660
ACGTCCTACG TC	CAGTGCCA	GGGGATCCCG	CAGGGCTCCA	TCCTCTCCAC	GCTGCTCTGC	720
AGCCTGTGCT AC	GCGGACAT	GGAGAACAAG	CTGTTTGCGG	GGATTGCGCG	GGACGGGCTG	780
CTCCTGCGTT TG	GGTGATGA	TTTCTTGTTG	GTGACACCTC	ACCTCACCCA	CGCGAAAACC	840
TTCTCAGGA CC	CTGGTCCG	AGGTGTCCCT	GAGTATGGCT	GCGTGGTGAA	CTTGCGGAAG	900
ACAGTGGTGA AC	TTCCCTGT	AGAAGACGAG	GCCCTGGGTG	GCACGGCTTT	TGTT	960
CCGGCCACG GC	CTATTCCC	CTGGTGCGGC	CTGCTGCTGG	ATACCCGGAC	CCTGGAGGTG	1020
CAGAGCGACT AC	TCCAGCTA	TGCCCCGACC	TCCATCAGAG	CCAGTCTCAC	CTTCAACCGC	1080
GGCTTCAAGG CT	TGGGAGGAA	CATGCGTCGC	AAACTCTTTG	GGGTCTTGCG	GCTGAAGTGT	1140
CACAGCCTGT TT	CTGGATTT	GCAGGTGAAC	AGCCTCCAGA	CGGTGTGCAC	CAACATCTAC	1200
AAGATCCTCC TG	CTGCAGGC	GTACAGGTTT	CACGCATGTG	TGCTGCAGCT	CCCATTTCAT	1260
CAGCAAGTTT GG	AAGAACCC	ACATTTTTTC	TGCGCGTCAT	CTCTGACACG	GCTCCCTCTG	1320
CTACTCCATC CT	TGAAAGCCA	AGAACGCAGG	GATGTCGCTG	GGGGCCAAGG	GCGCCGCCGG	1380
CCCTCTGCCC TC	CGAGGCCG	TGCAGTGGCT	GTGCCACCAA	GCATTCTGCT	TCAAGCTGAC	1440
TCGACACCGT GT	CACCTACG	TGCCACTCCT	GGGGTCACTC	AGGACAGCCC	AGACGCAGCT	1500
GAGTCGGAAG CT	CCCCGGGA	CGACGCTGAC	TGCCCTGGAG	GCCGCAGCCA	ACCCGGCACT	1560
GCCCTCAGAC TT	CAAGACCA	TCCTGGACTG	ATGGCCACCC	GCCACAGCC	AGGCCGAGAG	1620
CAGACACCAG CA	GCCCTGTC	ACGCCGGGCT	TATACGTCCC	AGGGAGGGAG	GGGCGGCCCA	1680
CACCCAGGCC TG	CACCGCTG	GGAGTCTGAG	GCCTGAGTGA	GTGTTTGGCC	GAGGCCTGCA	1740

```

TGTCGGGCTG AAGGCTGAGT GTCCGGCTGA GGCCTGAGCG AGTGTCCAGC CAAGGGCTGA      1800
GTGTCCAGCA CACCTGCGTT TTCACTTCCC CACAGGCTGG CGTTCGGTCC ACCCCAGGGC      1860
CAGCTTTTCC TCACCAGGAG CCCGGCTTCC ACTCCCCACA TAGGAATAGT CCATCCCCAG      1920
ATTGCGCATT GTTCACCCTT CGCCCTGCCT TCCTTTGCCT TCCACCCCCA CCATTCAGGT      1980
GGAGACCCTG AGAAGGACCC TGGGAGCTTT GGGAAATTTGG AGTGACCAAA GGTGTGCCCT      2040
GTACACAGGC GAGGACCCTG CACCTGGATG GGGGTCCCTG TGGGTCAAAT TGGGGGGAGG      2100
TGCTGTGGGA GTAAAATACT GAATATATGA GTTTTTCAGT TTTGGAAAAA AAAAAAAAAA      2160
AAAAAAAAA A                                                                2171

```

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 564 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..564
- (D) OTHER INFORMATION: /note= "63 kDa telomerase protein  
encoded by ORF of EcoRI-NotI insert of  
clone 712562"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```

Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr
1           5           10           15
Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val
20          25          30
Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu Lys Arg Val
35          40          45
Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala
50          55          60
Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp
65          70          75          80
Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly Ala Arg Thr
85          90          95
Phe Arg Arg Glu Lys Xaa Ala Glu Arg Leu Thr Ser Arg Val Lys Ala
100         105         110
Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu
115        120        125

```

Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala Trp Arg Thr  
 130 135 140  
 Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu Leu Tyr Phe  
 145 150 155 160  
 Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg  
 165 170 175  
 Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys  
 180 185 190  
 Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala Met Gly Thr Ser Ala  
 195 200 205  
 Arg Pro Ser Arg Ala Thr Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln  
 210 215 220  
 Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met  
 225 230 235 240  
 Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg  
 245 250 255  
 Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala Lys  
 260 265 270  
 Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys Val  
 275 280 285  
 Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu Ala  
 290 295 300  
 Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe Pro  
 305 310 315 320  
 Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser Asp  
 325 330 335  
 Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe Asn  
 340 345 350  
 Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly Val  
 355 360 365  
 Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn Ser  
 370 375 380  
 Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln Ala  
 385 390 395 400  
 Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln Val  
 405 410 415  
 Trp Lys Asn Pro His Phe Ser Cys Ala Ser Ser Leu Thr Arg Leu Pro  
 420 425 430  
 Leu Leu Leu His Pro Glu Ser Gln Glu Arg Arg Asp Val Ala Gly Gly  
 435 440 445  
 Gln Gly Arg Arg Arg Pro Ser Ala Leu Arg Gly Arg Ala Val Ala Val  
 450 455 460

Pro Pro Ser Ile Pro Ala Gln Ala Asp Ser Thr Pro Cys His Leu Arg  
 465 470 475 480

Ala Thr Pro Gly Val Thr Gln Asp Ser Pro Asp Ala Ala Glu Ser Glu  
 485 490 495

Ala Pro Gly Asp Asp Ala Asp Cys Pro Gly Gly Arg Ser Gln Pro Gly  
 500 505 510

Thr Ala Leu Arg Leu Gln Asp His Pro Gly Leu Met Ala Thr Arg Pro  
 515 520 525

Gln Pro Gly Arg Glu Gln Thr Pro Ala Ala Leu Ser Arg Arg Ala Tyr  
 530 535 540

Thr Ser Gln Gly Gly Arg Gly Gly Pro His Pro Gly Leu His Arg Trp  
 545 550 555 560

Glu Ser Glu Ala

(2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 50 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CCAGTGAGCA GAGTGACGAG GACTCGAGCT CAAGCTTTTT TTTTTTTTTT

50

(2) INFORMATION FOR SEQ ID NO:103:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

CCAGTGAGCA GAGTGACG

18

(2) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

GAGGACTCGA GCTCAAGC 18

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

CACTGATCCT TTCTTTTTCG TAAACGATAG GT 32

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

CATCAATCAA ATCTTCCATA TAGAAATGAC A 31

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION: 1
- (D) OTHER INFORMATION: /mod\_base= OTHER  
/note= "N = 5'-phosphorylated guanosine"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

NGGCCGTGTT GGCCTAGTTC TCTGCTC 27

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

GAGGAGGAGA AGAGCAGAGA ACTAGGCCAA CACGCCCC

38

(2) INFORMATION FOR SEQ ID NO:109:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

GTGTCATTTC TATATGGAAG ATTTGATTGA TG

32

(2) INFORMATION FOR SEQ ID NO:110:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

ACCTATCGTT TACGAAAAG AAAGGATCAG TG

32

(2) INFORMATION FOR SEQ ID NO:111:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GAGTGACATA ATATACGTGA

20

(2) INFORMATION FOR SEQ ID NO:112:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 amino acids
  - (B) TYPE: amino acid



- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Phe Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe  
1 5 10 15  
Tyr Arg Lys Ser Val Trp Ser Lys  
20

(2) INFORMATION FOR SEQ ID NO:113:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Arg Gln His Leu Lys Arg Val Gln Leu Arg Asp Val Ser Glu Ala Glu  
1 5 10 15  
Val Arg Gln His Arg Glu Ala  
20

(2) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg  
1 5 10 15  
Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu  
20 25

(2) INFORMATION FOR SEQ ID NO:115:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

Ala Lys Phe Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu  
1 5 10 15

Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr Phe Gln  
20 25

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly  
1 5 10 15

Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg Asp Val Ser  
20 25 30

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly  
1 5 10 15

Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val  
20 25

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

YARACHAARG GHATYCCHYA RGG

23

(2) INFORMATION FOR SEQ ID NO:119:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Gln Thr Lys Gly Ile Pro Gln Gly  
1 5

(2) INFORMATION FOR SEQ ID NO:120:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

NGTNATDARD ARRTARTCRT C

21

(2) INFORMATION FOR SEQ ID NO:121:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Asp Asp Tyr Leu Leu Ile Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:122:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Lys Gly Ile Pro Gln Gly Leu Cys Val Ser Ser Ile Leu Ser Ser Phe  
1 5 10 15  
Tyr Tyr Ala Thr Leu Glu Glu Ser Ser Leu Gly Phe Leu Arg Asp Glu  
20 25 30  
Ser Met Asn Pro Glu Asn Pro Asn Val Asn Leu Leu Met Arg Leu Thr  
35 40 45  
Asp Asp Tyr Leu Leu Ile Thr  
50 55

(2) INFORMATION FOR SEQ ID NO:123:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Ser Ile Leu Ser Ser Phe Leu Cys His Phe Tyr Met Glu Asp Leu Ile  
1 5 10 15  
Asp Glu Tyr Leu Ser Phe Thr Lys Lys Lys Gly Ser Val Leu Leu Arg  
20 25 30  
Val Val

(2) INFORMATION FOR SEQ ID NO:124:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 49 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Asp Gly Leu Phe Gln Gly Ser Ser Leu Ser Ala Pro Ile Val Asp Leu  
1 5 10 15  
Val Tyr Asp Asp Leu Leu Glu Phe Tyr Ser Glu Phe Lys Ala Ser Pro  
20 25 30  
Ser Gln Asp Thr Leu Ile Leu Lys Leu Ala Asp Asp Phe Leu Ile Ile  
35 40 45  
Ser

(2) INFORMATION FOR SEQ ID NO:125:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Gln Lys Val Gly Ile Pro Gln Gly  
1 5

(2) INFORMATION FOR SEQ ID NO:126:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

CAAAAAGTTG GTATCCCTCA GGG

23

(2) INFORMATION FOR SEQ ID NO:127:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 146 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

AGACCAAAGG AATTCCATCA GGCTCAATTC TGTCATCTTT TTTGTGTCAT TTCTATATGG 60  
AAGATTTGAT TGATGAATAC CTATCGTTTA CGAAAAAGAA AGGATCAGTG TTGTTACGAG 120  
TAGTCGACGA CTACCTCCTC ATCACC 146

(2) INFORMATION FOR SEQ ID NO:128:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 47 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Lys Gly Ile Pro Ser Gly Ser Ile Leu Ser Ser Phe Leu Cys His Phe  
1 5 10 15  
Tyr Met Glu Asp Leu Ile Asp Glu Tyr Leu Ser Phe Thr Lys Lys Lys  
20 25 30  
Gly Ser Val Leu Leu Arg Val Val Asp Asp Tyr Leu Leu Ile Thr  
35 40 45

(2) INFORMATION FOR SEQ ID NO:129:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

GACGATTTC TCCTTATAAC A

21

(2) INFORMATION FOR SEQ ID NO:130:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Asp Asp Phe Leu Phe Ile Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:131:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

AAAAAAAAAA AAAAAA

16

(2) INFORMATION FOR SEQ ID NO:132:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

TTTTTTTTTT TTTTTT

17

(2) INFORMATION FOR SEQ ID NO:133:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..35
- (D) OTHER INFORMATION: /note= "motif 0 peptide from Schizosaccharomyces pombe tez1p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

Trp	Leu	Tyr	Asn	Ser	Phe	Ile	Ile	Pro	Ile	Leu	Gln	Ser	Phe	Phe	Tyr
1				5				10						15	
Ile	Thr	Glu	Ser	Ser	Asp	Leu	Arg	Asn	Arg	Thr	Val	Tyr	Phe	Arg	Lys
			20					25					30		
Asp	Ile	Trp													
		35													

(2) INFORMATION FOR SEQ ID NO:134:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from Schizosaccharomyces pombe tez1p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

Ala	Val	Ile	Arg	Leu	Leu	Pro	Lys	Lys	Asn	Thr	Phe	Arg	Leu	Ile	Thr
1				5					10					15	

Asn Leu Arg Lys Arg Phe  
20

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..26
- (D) OTHER INFORMATION: /note= "motif 3(A) peptide from Schizosaccharomyces pombe tezlp"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Lys Lys Tyr Phe Val Arg Ile Asp Ile Lys Ser Cys Tyr Asp Arg Ile  
1 5 10 15

Lys Gln Asp Leu Met Phe Arg Ile Val Lys  
20 25

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..32
- (D) OTHER INFORMATION: /note= "motif 4(B') peptide from Schizosaccharomyces pombe tezlp"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

Tyr Leu Gln Lys Val Gly Ile Pro Gln Gly Ser Ile Leu Ser Ser Phe  
1 5 10 15

Leu Cys His Phe Tyr Met Glu Asp Leu Ile Asp Glu Tyr Leu Ser Phe  
20 25 30

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 amino acids



- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..49
- (D) OTHER INFORMATION: /note= "motif 5(C) and 6(D) peptide from Schizosaccharomyces pombe tezlp"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

```

Val Leu Leu Arg Val Val Asp Asp Phe Leu Phe Ile Thr Val Asn Lys
1           5           10           15
Lys Asp Ala Lys Lys Phe Leu Asn Leu Ser Leu Arg Gly Phe Glu Lys
                20           25           30
His Asn Phe Ser Thr Ser Leu Glu Lys Thr Val Ile Asn Phe Glu Asn
        35           40           45
Ser

```

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..34
- (D) OTHER INFORMATION: /note= "motif 0 peptide from Saccharomyces cerevisiae EST2p"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

```

Trp Leu Phe Arg Gln Leu Ile Pro Lys Ile Ile Gln Thr Phe Phe Tyr
1           5           10           15
Cys Thr Glu Ile Ser Ser Thr Val Thr Ile Val Tyr Phe Arg His Asp
        20           25           30
Thr Trp

```

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid

(ii) MOLECULE TYPE: peptide

(D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from  
Saccharomyces cerevisiae EST2p"

Ser Lys Met Arg Ile Ile Pro Lys Lys Ser Asn Asn Glu Phe Arg Ile  
1 5 10 15

Ile Ala Ile Pro Cys Arg Gly Ala Asp  
20 25

(D) TOPOLOGY: linear

(D) OTHER INFORMATION: /note= "motif 3(A) peptide from  
Saccharomyces cerevisiae EST2p"

Glu Leu Tyr Phe Met Lys Phe Asp Val Lys Ser Cys Tyr Asp Ser Ile  
1 5 10 15

Pro Arg Met Glu Cys Met Arg Ile Leu Lys  
20 25

(D) TOPOLOGY: linear

(A) NAME/KEY: Peptide

(D) OTHER INFORMATION: /note= "motif 4(B') peptide from  
Saccharomyces cerevisiae EST2p"

Tyr Ile Arg Glu Asp Gly Leu Phe Gln Gly Ser Ser Leu Ser Ala Pro  
1 5 10 15  
Ile Val Asp Leu Val Tyr Asp Asp Leu Leu Glu Phe Tyr Ser Glu Phe  
20 25 30

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 49 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..49  
(D) OTHER INFORMATION: /note= "motif 5(C) peptide from  
Saccharomyces cerevisiae EST2p"

Leu Ile Leu Lys Leu Ala Asp Asp Phe Leu Ile Ile Ser Thr Asp Gln  
1 5 10 15  
Gln Gln Val Ile Asn Ile Lys Lys Leu Ala Met Gly Gly Phe Gln Lys  
20 25 30  
Tyr Asn Ala Lys Ala Asn Arg Asp Lys Ile Leu Ala Val Ser Ser Gln  
35 40 45  
Ser

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(A) NAME/KEY: Peptide  
(B) LOCATION: 1..35  
(D) OTHER INFORMATION: /note= "motif 0 peptide from  
Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

Trp Ile Phe Glu Asp Leu Val Val Ser Leu Ile Arg Cys Phe Phe Tyr  
1                   5                   10                   15  
Val Thr Glu Gln Gln Lys Ser Tyr Ser Lys Thr Tyr Tyr Tyr Arg Lys  
                  20                   25                   30  
Asn Ile Trp  
          35

(2) INFORMATION FOR SEQ ID NO:144:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 23 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
    (A) NAME/KEY: Peptide  
    (B) LOCATION: 1..23  
    (D) OTHER INFORMATION: /note= "motif 1 and 2 peptide from  
                                  Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

Gly Lys Leu Arg Leu Ile Pro Lys Lys Thr Thr Phe Arg Pro Ile Met  
1                   5                   10                   15  
Thr Phe Asn Lys Lys Ile Val  
          20

(2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (ix) FEATURE:  
    (A) NAME/KEY: Peptide  
    (B) LOCATION: 1..26  
    (D) OTHER INFORMATION: /note= "motif 3(A) peptide from  
                                  Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr Asp Ser Val  
1                   5                   10                   15

Asn Arg Glu Lys Leu Ser Thr Phe Leu Lys  
20 25

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..32
- (D) OTHER INFORMATION: /note= "motif 4(B') peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

Tyr Lys Gln Thr Lys Gly Ile Pro Gln Gly Leu Cys Val Ser Ser Ile  
1 5 10 15  
Leu Ser Ser Phe Tyr Tyr Ala Thr Leu Glu Glu Ser Ser Leu Gly Phe  
20 25 30

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..49
- (D) OTHER INFORMATION: /note= "motif 5(C) and 6(D) peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu Ile Thr Thr Gln Glu  
1 5 10 15  
Asn Asn Ala Val Leu Phe Ile Glu Lys Leu Ile Asn Val Ser Arg Glu  
20 25 30  
Asn Gly Phe Lys Phe Asn Met Lys Lys Leu Gln Thr Ser Phe Pro Leu  
35 40 45  
Ser

(2) INFORMATION FOR SEQ ID NO:148:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..22
  - (D) OTHER INFORMATION: /note= "motif 1 peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

Leu Val Val Ser Leu Ile Arg Cys Phe Phe Tyr Val Thr Glu Gln Gln  
1                   5                   10                   15  
Lys Ser Tyr Ser Lys Thr  
                  20

(2) INFORMATION FOR SEQ ID NO:149:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..30
  - (D) OTHER INFORMATION: /note= "motif 0 peptide from Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

Lys Ser Leu Gly Phe Ala Pro Gly Lys Leu Arg Leu Ile Pro Lys Lys  
1                   5                   10                   15  
Thr Thr Phe Arg Pro Ile Met Thr Phe Asn Lys Lys Ile Val  
                  20                   25                   30

(2) INFORMATION FOR SEQ ID NO:150:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

· (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 1..27  
(D) OTHER INFORMATION: /note= "motif A peptide from  
Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

Pro Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr Asp Ser  
1 5 10 15

Val Asn Arg Glu Lys Leu Ser Thr Phe Leu Lys  
20 25

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 1..27  
(D) OTHER INFORMATION: /note= "motif B peptide from  
Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Asn Gly Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro Gln Gly Leu Cys  
1 5 10 15

Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala  
20 25

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 22  
(D) OTHER INFORMATION: /note= "motif C peptide from  
Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

Pro Asn Val Asn Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu Leu Ile  
1                      5                      10                      15  
Thr Thr Gln Glu Asn Asn  
                    20

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 15 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
    (A) NAME/KEY: Peptide  
    (B) LOCATION: 1..15  
    (D) OTHER INFORMATION: /note= "motif D peptide from  
                                    Euplotes aediculatus p123"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

Asn Val Ser Arg Glu Asn Gly Phe Lys Phe Asn Met Lys Lys Leu  
1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
    (A) NAME/KEY: Peptide  
    (B) LOCATION: 1..22  
    (D) OTHER INFORMATION: /note= "motif 1 peptide from  
                                    Schizosaccharomyces pombe tez1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

Phe Ile Ile Pro Ile Leu Gln Ser Phe Phe Tyr Ile Thr Glu Ser Ser  
1                      5                      10                      15  
Asp Leu Arg Asn Arg Thr  
                    20

(2) INFORMATION FOR SEQ ID NO:155:



- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:
  - (A) NAME/KEY: Peptide
  - (B) LOCATION: 1..30
  - (D) OTHER INFORMATION: /note= "motif 0 peptide from Schizosaccharomyces pombe tez1"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:
 

Gln	Lys	Thr	Thr	Leu	Pro	Pro	Ala	Val	Ile	Arg	Leu	Leu	Pro	Lys	Lys
1				5					10					15	
Asn	Thr	Phe	Arg	Leu	Ile	Thr	Asn	Leu	Arg	Lys	Arg	Phe	Leu		
			20				25						30		
- (2) INFORMATION FOR SEQ ID NO:156:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 27 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS:
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (ix) FEATURE:
    - (A) NAME/KEY: Peptide
    - (B) LOCATION: 1..27
    - (D) OTHER INFORMATION: /note= "motif A peptide from Schizosaccharomyces pombe tez1"
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:
 

Arg	Lys	Lys	Tyr	Phe	Val	Arg	Ile	Asp	Ile	Lys	Ser	Cys	Tyr	Asp	Arg
1				5					10					15	
Ile	Lys	Gln	Asp	Leu	Met	Phe	Arg	Ile	Val	Lys					
			20					25							
- (2) INFORMATION FOR SEQ ID NO:157:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 27 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS:
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..27
- (D) OTHER INFORMATION: /note= "motif B peptide from Schizosaccharomyces pombe tez1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

Gly Asn Ser Gln Tyr Leu Gln Lys Val Gly Ile Pro Gln Gly Ser Ile  
1                      5                      10                      15  
Leu Ser Ser Phe Leu Cys His Phe Tyr Met Glu  
                    20                      25

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "motif C peptide from Schizosaccharomyces pombe tez1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

Lys Lys Gly Ser Val Leu Leu Arg Val Val Asp Asp Phe Leu Phe Ile  
1                      5                      10                      15  
Thr Val Asn Lys Lys Asp  
                    20

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..15
- (D) OTHER INFORMATION: /note= "motif D peptide from Schizosaccharomyces pombe tez1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

Leu Asn Leu Ser Leu Arg Gly Phe Glu Lys His Asn Phe Ser Thr  
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "motif 1 peptide from  
Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

Leu Ile Pro Lys Ile Ile Gln Thr Phe Phe Tyr Cys Thr Glu Ile Ser  
 1 5 10 15  
 Ser Thr Val Thr Ile Val  
 20

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..32
- (D) OTHER INFORMATION: /note= "motif 0 peptide from  
Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

Thr Leu Ser Asn Phe Asn His Ser Lys Met Arg Ile Ile Pro Lys Lys  
 1 5 10 15  
 Ser Asn Asn Glu Phe Arg Ile Ile Ala Ile Pro Cys Arg Gly Ala Asp  
 20 25 30

(2) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid

- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..27
- (D) OTHER INFORMATION: /note= "motif A peptide from  
Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

Pro	Glu	Leu	Tyr	Phe	Met	Lys	Phe	Asp	Val	Lys	Ser	Cys	Tyr	Asp	Ser
1				5				10						15	
Ile	Pro	Arg	Met	Glu	Cys	Met	Arg	Ile	Leu	Lys					
			20					25							

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..27
- (D) OTHER INFORMATION: /note= "motif B peptide from  
Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

Glu	Asp	Lys	Cys	Tyr	Ile	Arg	Glu	Asp	Gly	Leu	Phe	Gln	Gly	Ser	Ser
1				5					10					15	
Leu	Ser	Ala	Pro	Ile	Val	Asp	Leu	Val	Tyr	Asp					
			20					25							

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide

(B) LOCATION: 1..22  
(D) OTHER INFORMATION: /note= "motif C peptide from  
Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

Ser Gln Asp Thr Leu Ile Leu Lys Leu Ala Asp Asp Phe Leu Ile Ile  
1 5 10 15  
Ser Thr Asp Gln Gln Gln  
20

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..15  
(D) OTHER INFORMATION: /note= "motif D peptide from  
Saccharomyces cerevisiae EST2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

Lys Lys Leu Ala Met Gly Gly Phe Gln Lys Tyr Asn Ala Lys Ala  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:  
(A) NAME/KEY: Peptide  
(B) LOCATION: 1..22  
(D) OTHER INFORMATION: /note= "motif 1 peptide from human  
telomerase core protein 1 (TCP1)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

Tyr Val Val Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr  
1 5 10 15  
Phe Gln Lys Asn Arg Leu  
20

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..30
- (D) OTHER INFORMATION: /note= "motif 0 peptide from human telomerase core protein 1 (TCP1)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

Ala	Arg	Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu	Arg	Phe	Ile	Pro	Lys	Pro
1				5					10					15	
Asp	Gly	Leu	Arg	Pro	Ile	Val	Asn	Met	Asp	Tyr	Val	Val	Gly		
			20					25					30		

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..27
- (D) OTHER INFORMATION: /note= "motif A peptide from human telomerase core protein 1 (TCP1)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

Pro	Glu	Leu	Tyr	Phe	Val	Lys	Val	Asp	Val	Thr	Gly	Ala	Tyr	Asp	Thr
1				5					10					15	
Ile	Pro	Gln	Asp	Arg	Leu	Thr	Glu	Val	Ile	Ala					
			20					25							

(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Protein

(B) LOCATION: 1..27

(D) OTHER INFORMATION: /note= "motif B peptide from human telomerase core protein 1 (TCP1)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

Arg Ala Thr Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile  
1 5 10 15

Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly  
20 25

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..22

(D) OTHER INFORMATION: /note= "motif C peptide from human telomerase core protein 1 (TCP1)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

Arg Arg Asp Gly Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val  
1 5 10 15

Thr Pro His Leu Thr His  
20

(2) INFORMATION FOR SEQ ID NO:171:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 1..15

(D) OTHER INFORMATION: /note= "motif D peptide from human telomerase core protein 1 (TCP1)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys Val Val  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

Phe Phe Tyr Val Thr Glu  
1 5

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4029 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY:  
(B) LOCATION: 1..4029  
(D) OTHER INFORMATION: /note= "preliminary sequence for  
human TRT cDNA insert of  
plasmid pGRN121"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC CCGCGATGCC	60
GCGCGCTCCC CGCTGCCGAG CCGTGCGCTC CCTGCTGCGC AGCCACTACC GCGAGGTGCT	120
GCCGCTGGCC ACGTTCGTGC GGCGCCTGGG GCCCCAGGGC TGGCGGCTGG TGCAGCGCGG	180
GGACCCGGCG GCTTTCCGCG CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN	240
ANGGCNGCCC CCCGCCGCC CCTCCTTCCG CCAGGTGTCC TGCCTGAANG ANCTGGTGGC	300
CCGAGTGCTG CANANGCTGT GCGANCGCGG CGCGAANAAC GTGCTGGCCT TCGGCTTCGC	360
GCTGCTGGAC GGGGCCCGCG GGGGCCCCC CGAGGCCTTC ACCACCAGCG TGCGCAGCTA	420
CCTGCCCAAC ACGGTGACCG ACGCACTGCG GGGGAGCGGG GCGTGGGGGC TGCTGCTGCG	480
CCGCGTGGGC GACGACGTGC TGGTTCACCT GCTGGCACGC TCGCGGNTNT TTGTGCTGGT	540



GGNTCCCAGC	TGCGCCTACC	ANGTGTGCGG	GCCGCCGCTG	TACCAGCTCG	GCGCTGCNAC	600
TCAGGCCCCG	CCCCCGCCAC	ACGCTANTGG	ACCCGAANGC	GTCTGGGATC	CAACGGGCCCT	660
GGAACCATAG	CGTCAGGGAG	GCCGGGGTCC	CCCTGGGCTG	CCAGCCCCGG	GTGCGAGGAG	720
GCGCGGGGGC	AGTGCCAGCC	GAAGTCTGCC	GTTGCCCAAG	AGGCCCAGGC	GTGGCGCTGC	780
CCCTGAGCCG	GAGCGGACGC	CCGTTGGGCA	GGGGTCCTGG	GCCCACCCGG	GCAGGACGCC	840
TGGACCGAGT	GACCGTGTTT	TCTGTGTGGT	GTCACCTGCC	AGACCCGCCG	AAGAAGCCAC	900
CTCTTTGGAG	GGTGCCTCT	CTGGCACGCG	CCACTCCAC	CCATCCGTGG	GCCGCCAGCA	960
CCACGCGGGC	CCCCCATCCA	CATCGCGGCC	ACCACGTCCT	GGGACACGCC	TTGTCCCCCG	1020
GTGTACGCCG	AGACCAAGCA	CTTCCTCTAC	TCCTCAGGCG	ACAAGNACAC	TGCGNCCCTC	1080
CTTCCTACTC	AATATATCTG	AGGCCCAGCC	TGACTGGCGT	TCGGGAGGTT	CGTGGAGACA	1140
NTCTTTCTGG	TTCCAGGCCT	TGGATGCCAG	GATTCCCCGC	AGGTTGCCCC	GCCTGCCCCA	1200
GCGNTACTGG	CAAATGCGGC	CCCTGTTTCT	GGAGCTGCTT	GGGAACCACG	CGCAGTGCCC	1260
CTACGGGGTG	TTCTCAAGA	CGCACTGCCC	GCTGCGAGCT	GCGGTCACCC	CAGCAGCCGG	1320
TGTCTGTGCC	CGGGAGAAGC	CCCAGGGCTC	TGTGGCGGCC	CCCGAGGAGG	AGGAACACAG	1380
ACCCCCGTCTG	CCTGGTGCAG	CTGCTCCGCC	AGCACAGCAG	CCCCTGGCAG	GTGTACGGCT	1440
TCGTGCGGGC	CTGCCTGCGC	CGGCTGGTGC	CCCCAGGCCT	CTGGGGCTCC	AGGCACAACG	1500
AACGCCGCTT	CCTCAGGAAC	ACCAAGAAGT	TCATCTCCCT	GGGGAAGCAT	GCCAAGCTCT	1560
CGCTGCAGGA	GCTGACGTGG	AAGATGAGCG	TGCGGGACTG	CGCTTGGCTG	CGCAGGAGCC	1620
CAGGGGTTGG	CTGTGTTCCG	GCCGCAGAGC	ACCGTCTGCG	TGAGGAGATC	CTGGCCAAGT	1680
TCCTGCACTG	GCTGATGAGT	GTGTACGTCG	TCGAGCTGCT	CAGGTCTTTC	TTTTATGTCA	1740
CGGAGACCAC	GTTTCAAAAG	AACAGGCTCT	TTTTCTACCG	GAAGAGTGTC	TGGAGCAAGT	1800
TGCAAAGCAT	TGGAATCAGA	CAGCACTTGA	AGAGGGTGCA	GCTGCGGGAG	CTGTGCGAAG	1860
CAGAGGTCAG	GCAGCATCGG	GAAGCCAGGC	CCGCCCTGCT	GACGTCCAGA	CTCCGCTTCA	1920
TCCCCAAGCC	TGACGGGCTG	CGGCCGATTG	TGAACATGGA	CTACGTCGTG	GGAGCCAGAA	1980
CGTTCCGCAG	AGAAAAGAGG	GCCGAGCGTC	TCACCTCGAG	GGTGAAGGCA	CTGTTCAGCG	2040
TGCTCAACTA	CGAGCGGGCG	CGGCGCCCCG	GCCTCCTGGG	CGCCTCTGTG	CTGGGCCTGG	2100
ACGATATCCA	CAGGGCCTGG	CGCACCTTCG	TGCTGCGTGT	GCGGGCCCAG	GACCCGCCGC	2160
CTGAGCTGTA	CTTTGTCAAG	GTGGATGTGA	CGGGCGCGTA	CGACACCATC	CCCCAGGACA	2220
GGCTCACGGA	GGTCATCGCC	AGCATCATCA	AACCCAGAA	CACGTACTGC	GTGCGTCGGT	2280
ATGCCGTGGT	CCAGAAGGCC	GCCCATGGGC	ACGTCCGCAA	GGCCTTCAAG	AGCCACGTCT	2340
CTACCTTGAC	AGACCTCCAG	CCGTACATGC	GACAGTTCGT	GGCTCACCTG	CAGGANAACA	2400

GECCGCTGAG	GGATGCCGTC	GTCATCGAGC	AGAGCTCCTC	CCTGAATGAG	GCCAGCAGTG	2460
GCCTCTTCGA	CGTCTTCCTA	CGCTTCATGT	GCCACCACGC	CGTGCGCATC	AGGGGCAAGT	2520
CCTACGTCCA	GTGCCAGGGG	ATCCCGCAGG	GCTCCATCCT	CTCCACGCTG	CTCTGCAGCC	2580
TGTGCTACGG	CGACATGGAG	AACAAGCTGT	TTGCGGGGAT	TCGGCGGGAC	GGGCTGCTCC	2640
TGCGTTTGGT	GGATGATTTT	TTGTTGGTGA	CACCTCACCT	CACCCACGCG	AAAACCTTCC	2700
TCAGGACCCT	GGTCCGAGGT	GTCCCTGAGT	ATGGCTGCGT	GGTGAACTTG	CGGAAGACAG	2760
TGGTGAACTT	CCCTGTAGAA	GACGAGGCCC	TGGGTGGCAC	GGCTTTTGTT	CAGATGCCGG	2820
CCCACGGCCT	ATTCCCCTGG	TGCGGCCTGC	TGCTGGATAC	CCGGACCCTG	GAGGTGCAGA	2880
GCGACTACTC	CAGCTATGCC	CGGACCTCCA	TCAGAGCCAG	TCTCACCTTC	AACCGCGGCT	2940
TCAAGGCTGG	GAGGAACATG	CGTCGCAAAC	TCTTTGGGGT	CTTGCGGCTG	AAGTGTCA	3000
GCCTGTTTCT	GGATTTGCAG	GTGAACAGCC	TCCAGACGGT	GTGCACCAAC	ATCTACAAGA	3060
TCCTCCTGCT	GCAGGCGTAC	AGGTTTCACG	CATGTGTGCT	GCAGCTCCCA	TTTCATCAGC	3120
AAGTTTGGA	GAACCCACA	TTTTTCCTGC	GCGTCATCTC	TGACACGGCC	TCCCTCTGCT	3180
ACTCCATCCT	GAAAGCCAAG	AACGCAGGGA	TGTCGCTGGG	GGCCAAGGGC	GCCGCCGGCC	3240
CTCTGCCCTC	CGAGGCCGTG	CAGTGGCTGT	GCCACCAAGC	ATTCCTGCTC	AAGCTGACTC	3300
GACACCGTGT	CACCTACGTG	CCACTCCTGG	GGTCACTCAG	GACAGCCCAG	ACGCAGCTGA	3360
GTCGGAAGCT	CCCGGGGACG	ACGCTGACTG	CCCTGGAGGC	CGCAGCCAAC	CCGGCACTGC	3420
CCTCAGACTT	CAAGACCATC	CTGGACTGAT	GGCCACCCGC	CCACAGCCAG	GCCGAGAGCA	3480
GACACCAGCA	GCCCTGTCAC	GCCGGGCTCT	ACGTCCCAGG	GAGGGAGGGG	CGGCCACAC	3540
CCAGGCCCCG	ACCGCTGGGA	GTCTGAGGCC	TGAGTGAGTG	TTTGCCGAG	GCCTGCATGT	3600
CCGGCTGAAG	GCTGAGTGTC	CGGCTGAGGC	CTGAGCGAGT	GTCCAGCCAA	GGGCTGAGTG	3660
TCCAGCACAC	CTGCCGTCTT	CACTTCCCCA	CAGGCTGGCG	CTCGGCTCCA	CCCCAGGGCC	3720
AGCTTTTCCT	CACCAGGAGC	CCGGCTTCCA	CTCCCCACAT	AGGAATAGTC	CATCCCCAGA	3780
TTGCCATTG	TTCACCCCTC	GCCCTGCCCT	CCTTTGCCTT	CCACCCCCAC	CATCCAGGTG	3840
GAGACCCTGA	GAAGGACCCT	GGGAGCTCTG	GGAATTTGGA	GTGACCAAAG	GTGTGCCCTG	3900
TACACAGGCG	AGGACCCTGC	ACCTGGATGG	GGGTCCCTGT	GGGTCAAATT	GGGGGGAGGT	3960
GCTGTGGGAG	TAAAATACTG	AATATATGAG	TTTTTCAGTT	TTGAAAAAAA	AAAAAAAAAA	4020
AAAAAAAAAA						4029

(2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 261 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

Ala	Ala	Leu	Arg	Pro	Ala	Ala	His	Val	Gly	Ser	Pro	Gly	Pro	Gly	His
1				5					10					15	
Pro	Arg	Asp	Ala	Ala	Arg	Ser	Pro	Leu	Pro	Ser	Arg	Ala	Leu	Pro	Ala
			20					25					30		
Ala	Gln	Pro	Leu	Pro	Arg	Gly	Ala	Ala	Ala	Gly	His	Val	Arg	Ala	Ala
		35				40					45				
Pro	Gly	Ala	Pro	Gly	Leu	Ala	Ala	Gly	Ala	Ala	Arg	Gly	Pro	Gly	Gly
	50					55					60				
Phe	Pro	Arg	Xaa	Gly	Gly	Pro	Xaa	Xaa	Gly	Val	Arg	Ala	Leu	Gly	Xaa
65				70					75					80	
Xaa	Ala	Ala	Pro	Arg	Arg	Pro	Leu	Leu	Pro	Pro	Gly	Val	Leu	Pro	Glu
				85					90					95	
Xaa	Xaa	Gly	Gly	Pro	Ser	Ala	Ala	Xaa	Ala	Val	Arg	Xaa	Arg	Arg	Glu
			100					105					110		
Xaa	Arg	Ala	Gly	Leu	Arg	Leu	Arg	Ala	Ala	Gly	Arg	Gly	Pro	Arg	Gly
	115					120						125			
Pro	Pro	Arg	Gly	Leu	His	His	Gln	Arg	Ala	Gln	Leu	Pro	Ala	Gln	His
	130					135					140				
Gly	Asp	Arg	Arg	Thr	Ala	Gly	Glu	Arg	Gly	Val	Gly	Ala	Ala	Ala	Ala
145					150					155					160
Pro	Arg	Gly	Arg	Arg	Arg	Ala	Gly	Ser	Pro	Ala	Gly	Thr	Leu	Arg	Xaa
				165					170					175	
Xaa	Cys	Ala	Gly	Gly	Ser	Gln	Leu	Arg	Leu	Pro	Xaa	Val	Arg	Ala	Ala
			180					185					190		
Ala	Val	Pro	Ala	Arg	Arg	Cys	Xaa	Ser	Gly	Pro	Ala	Pro	Ala	Thr	Arg
		195					200					205			
Xaa	Trp	Thr	Arg	Xaa	Arg	Leu	Gly	Ser	Asn	Gly	Pro	Gly	Thr	Ile	Ala
	210					215					220				
Ser	Gly	Arg	Pro	Gly	Ser	Pro	Trp	Ala	Ala	Ser	Pro	Gly	Cys	Glu	Glu
225					230					235				240	
Ala	Arg	Gly	Gln	Cys	Gln	Pro	Lys	Ser	Ala	Val	Ala	Gln	Glu	Ala	Gln
				245					250					255	
Ala	Trp	Arg	Cys	Pro											
				260											

(2) INFORMATION FOR SEQ ID NO:175:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

Ala Gly Ala Asp Ala Arg Trp Ala Gly Val Leu Gly Pro Pro Gly Gln  
 1                      5                      10                      15  
 Asp Ala Trp Thr Glu  
                     20

(2) INFORMATION FOR SEQ ID NO:176:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 82 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

Pro Trp Phe Leu Cys Gly Val Thr Cys Gln Thr Arg Arg Arg Ser His  
 1                      5                      10                      15  
 Leu Phe Gly Gly Cys Ala Leu Trp His Ala Pro Leu Pro Pro Ile Arg  
                     20                      25                      30  
 Gly Pro Pro Ala Pro Arg Gly Pro Pro Ile His Ile Ala Ala Thr Thr  
                     35                      40                      45  
 Ser Trp Asp Thr Pro Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe  
                     50                      55                      60  
 Leu Tyr Ser Ser Gly Asp Lys Xaa Thr Ala Xaa Leu Leu Pro Thr Gln  
 65                      70                      75                      80  
 Tyr Ile

(2) INFORMATION FOR SEQ ID NO:177:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 153 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

Leu Ala Phe Gly Arg Phe Val Glu Thr Xaa Phe Leu Val Pro Gly Leu  
 1 5 10 15  
 Gly Cys Gln Asp Ser Pro Gln Val Ala Pro Pro Ala Pro Ala Xaa Leu  
 20 25 30  
 Ala Asn Ala Ala Pro Val Ser Gly Ala Ala Trp Glu Pro Arg Ala Val  
 35 40 45  
 Pro Leu Arg Gly Val Pro Gln Asp Ala Leu Pro Ala Ala Ser Cys Gly  
 50 55 60  
 His Pro Ser Ser Arg Cys Leu Cys Pro Gly Glu Ala Pro Gly Leu Cys  
 65 70 75 80  
 Gly Gly Pro Arg Gly Gly Gly Thr Gln Thr Pro Val Ala Trp Cys Ser  
 85 90 95  
 Cys Ser Ala Ser Thr Ala Ala Pro Gly Arg Cys Thr Ala Ser Cys Gly  
 100 105 110  
 Pro Ala Cys Ala Gly Trp Cys Pro Gln Ala Ser Gly Ala Pro Gly Thr  
 115 120 125  
 Thr Asn Ala Ala Ser Ser Gly Thr Pro Arg Ser Ser Ser Pro Trp Gly  
 130 135 140  
 Ser Met Pro Ser Ser Arg Cys Arg Ser  
 145 150

(2) INFORMATION FOR SEQ ID NO:178:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

Ala Cys Gly Thr Ala Leu Gly Cys Ala Gly Ala Gln Gly Leu Ala Val  
 1 5 10 15  
 Phe Arg Pro Gln Ser Thr Val Cys Val Arg Arg Ser Trp Pro Ser Ser  
 20 25 30  
 Cys Thr Gly  
 35

(2) INFORMATION FOR SEQ ID NO:179:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

Val Cys Thr Ser Ser Ser Cys Ser Gly Leu Ser Phe Met Ser Arg Arg  
1 5 10 15  
Pro Arg Phe Lys Arg Thr Gly Ser Phe Ser Thr Gly Arg Val Ser Gly  
20 25 30  
Ala Ser Cys Lys Ala Leu Glu Ser Asp Ser Thr  
35 40

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

Arg Gly Cys Ser Cys Gly Ser Cys Arg Lys Gln Arg Ser Gly Ser Ile  
1 5 10 15  
Gly Lys Pro Gly Pro Pro Cys  
20

(2) INFORMATION FOR SEQ ID NO:181:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

Arg Pro Asp Ser Ala Ser Ser Pro Ser Leu Thr Gly Cys Gly Arg Leu  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

Thr Trp Thr Thr Ser Trp Glu Pro Glu Arg Ser Ala Glu Lys Arg Gly  
1 5 10 15  
Pro Ser Val Ser Pro Arg Gly  
20

(2) INFORMATION FOR SEQ ID NO:183:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 54 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

Arg His Cys Ser Ala Cys Ser Thr Thr Ser Gly Arg Gly Ala Pro Ala  
1 5 10 15  
Ser Trp Ala Pro Leu Cys Trp Ala Trp Thr Ile Ser Thr Gly Pro Gly  
20 25 30  
Ala Pro Ser Cys Cys Val Cys Gly Pro Arg Thr Arg Arg Leu Ser Cys  
35 40 45  
Thr Leu Ser Arg Trp Met  
50

(2) INFORMATION FOR SEQ ID NO:184:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 52 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

Arg Ala Arg Thr Thr Pro Ser Pro Arg Thr Gly Ser Arg Arg Ser Ser  
1 5 10 15  
Pro Ala Ser Ser Asn Pro Arg Thr Arg Thr Ala Cys Val Gly Met Pro  
20 25 30  
Trp Ser Arg Arg Pro Pro Met Gly Thr Ser Ala Arg Pro Ser Arg Ala  
35 40 45  
Thr Ser Leu Pro  
50

(2) INFORMATION FOR SEQ ID NO:185:

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

Gln Thr Ser Ser Arg Thr Cys Asp Ser Ser Trp Leu Thr Cys Arg Xaa  
1 5 10 15

Thr Ala Arg

(2) INFORMATION FOR SEQ ID NO:186:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

Gly Met Pro Ser Ser Ser Ser Arg Ala Pro Pro  
1 5 10

(2) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 74 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

Met Arg Pro Ala Val Ala Ser Ser Thr Ser Ser Tyr Ala Ser Cys Ala  
1 5 10 15

Thr Thr Pro Cys Ala Ser Gly Ala Ser Pro Thr Ser Ser Ala Arg Gly  
20 25 30

Ser Arg Arg Ala Pro Ser Ser Pro Arg Cys Ser Ala Ala Cys Ala Thr  
35 40 45

Ala Thr Trp Arg Thr Ser Cys Leu Arg Gly Phe Gly Gly Thr Gly Cys  
50 55 60

Ser Cys Val Trp Trp Met Ile Ser Cys Trp  
65 70



(2) INFORMATION FOR SEQ ID NO:188:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 24 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

His Leu Thr Ser Pro Thr Arg Lys Pro Ser Ser Gly Pro Trp Ser Glu  
1                    5                    10                    15  
Val Ser Leu Ser Met Ala Ala Trp  
                    20

(2) INFORMATION FOR SEQ ID NO:189:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 6 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

Thr Cys Gly Arg Gln Trp  
1                    5

(2) INFORMATION FOR SEQ ID NO:190:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 70 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

Lys Thr Arg Pro Trp Val Ala Arg Leu Leu Phe Arg Cys Arg Pro Thr  
1                    5                    10                    15  
Ala Tyr Ser Pro Gly Ala Ala Cys Cys Trp Ile Pro Gly Pro Trp Arg  
                    20                    25                    30  
Cys Arg Ala Thr Thr Pro Ala Met Pro Gly Pro Pro Ser Glu Pro Val  
                    35                    40                    45  
Ser Pro Ser Thr Ala Ala Ser Arg Leu Gly Gly Thr Cys Val Ala Asn  
                    50                    55                    60

Ser Leu Gly Ser Cys Gly  
65 70

(2) INFORMATION FOR SEQ ID NO:191:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

Ser Val Thr Ala Cys Phe Trp Ile Cys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:192:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

Thr Ala Ser Arg Arg Cys Ala Pro Thr Ser Thr Arg Ser Ser Cys Cys  
1 5 10 15  
Arg Arg Thr Gly Phe Thr His Val Cys Cys Ser Ser His Phe Ile Ser  
20 25 30  
Lys Phe Gly Arg Thr Pro His Phe Ser Cys Ala Ser Ser Leu Thr Arg  
35 40 45  
Pro Pro Ser Ala Thr Pro Ser  
50 55

(2) INFORMATION FOR SEQ ID NO:193:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

Lys Pro Arg Thr Gln Gly Cys Arg Trp Gly Pro Arg Ala Pro Pro Ala  
1 5 10 15

Leu Cys Pro Pro Arg Pro Cys Ser Gly Cys Ala Thr Lys His Ser Cys  
 20 25 30

Ser Ser

(2) INFORMATION FOR SEQ ID NO:194:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

Leu Asp Thr Val Ser Pro Thr Cys His Ser Trp Gly His Ser Gly Gln  
 1 5 10 15

Pro Arg Arg Ser  
 20

(2) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

Val Gly Ser Ser Arg Gly Arg Arg  
 1 5

(2) INFORMATION FOR SEQ ID NO:196:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 61 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

Leu Pro Trp Arg Pro Gln Pro Thr Arg His Cys Pro Gln Thr Ser Arg  
 1 5 10 15

Pro Ser Trp Thr Asp Gly His Pro Pro Thr Ala Arg Pro Arg Ala Asp  
 20 25 30

Thr Ser Ser Pro Val Thr Pro Gly Ser Thr Ser Gln Gly Gly Arg Gly  
 35 40 45

Gly Pro His Pro Gly Pro His Arg Trp Glu Ser Glu Ala  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:197:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 13 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

Val Ser Val Trp Pro Arg Pro Ala Cys Pro Ala Glu Gly  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:198:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

Gly Leu Ser Glu Cys Pro Ala Lys Gly  
 1 5

(2) INFORMATION FOR SEQ ID NO:199:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

Val Ser Ser Thr Pro Ala Val Phe Thr Ser Pro Gln Ala Gly Ala Arg  
 1 5 10 15

Leu His Pro Arg Ala Ser Phe Ser Ser Pro Gly Ala Arg Leu Pro Leu  
 20 25 30

Pro Thr

(2) INFORMATION FOR SEQ ID NO:200:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 70 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

```
Ser Ile Pro Arg Phe Ala Ile Val His Pro Ser Pro Cys Pro Pro Leu
1           5           10           15

Pro Ser Thr Pro Thr Ile Gln Val Glu Thr Leu Arg Arg Thr Leu Gly
          20           25           30

Ala Leu Gly Ile Trp Ser Asp Gln Arg Cys Ala Leu Tyr Thr Gly Glu
          35           40           45

Asp Pro Ala Pro Gly Trp Gly Ser Leu Trp Val Lys Leu Gly Gly Gly
          50           55           60

Ala Val Gly Val Lys Tyr
65           70
```

(2) INFORMATION FOR SEQ ID NO:201:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 16 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

```
Ile Tyr Glu Phe Phe Ser Phe Glu Lys Lys Lys Lys Lys Lys Lys
1           5           10           15
```

(2) INFORMATION FOR SEQ ID NO:202:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 222 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

```
Gln Arg Cys Val Leu Leu Arg Thr Trp Glu Ala Leu Ala Pro Ala Thr
1           5           10           15
```

Pro Ala Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu  
                     20                    25                    30  
 Arg Ser His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg  
                     35                    40                    45  
 Leu Gly Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala  
                     50                    55                    60  
 Phe Arg Ala Xaa Val Ala Xaa Cys Xaa Val Cys Val Pro Trp Xaa Xaa  
                     65                    70                    75                    80  
 Xaa Xaa Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Xaa  
                     85                    90                    95  
 Xaa Leu Val Ala Arg Val Leu Xaa Xaa Leu Cys Xaa Arg Gly Ala Xaa  
                     100                    105                    110  
 Asn Val Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly  
                     115                    120                    125  
 Pro Pro Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr  
                     130                    135                    140  
 Val Thr Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg  
                     145                    150                    155                    160  
 Arg Val Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Xaa  
                     165                    170                    175  
 Phe Val Leu Val Xaa Pro Ser Cys Ala Tyr Xaa Val Cys Gly Pro Pro  
                     180                    185                    190  
 Leu Tyr Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala  
                     195                    200                    205  
 Xaa Gly Pro Glu Xaa Val Trp Asp Pro Thr Gly Leu Glu Pro  
                     210                    215                    220

(2) INFORMATION FOR SEQ ID NO:203:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 330 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

Arg Gln Gly Gly Arg Gly Pro Pro Gly Leu Pro Ala Pro Gly Ala Arg  
 1                    5                    10                    15  
 Arg Arg Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro  
                     20                    25                    30  
 Arg Arg Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly  
                     35                    40                    45

Ser Trp Ala His Pro Gly Arg Thr Pro Gly Pro Ser Asp Arg Gly Phe  
 50 55 60  
 Cys Val Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu  
 65 70 75 80  
 Gly Ala Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln  
 85 90 95  
 His His Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Gly Thr  
 100 105 110  
 Arg Leu Val Pro Arg Cys Thr Pro Arg Pro Ser Thr Ser Ser Thr Pro  
 115 120 125  
 Gln Ala Thr Xaa Thr Leu Arg Pro Ser Phe Leu Leu Asn Ile Ser Glu  
 130 135 140  
 Ala Gln Pro Asp Trp Arg Ser Gly Gly Ser Trp Arg Xaa Ser Phe Trp  
 145 150 155 160  
 Phe Gln Ala Leu Asp Ala Arg Ile Pro Arg Arg Leu Pro Arg Leu Pro  
 165 170 175  
 Gln Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn  
 180 185 190  
 His Ala Gln Cys Pro Tyr Gly Val Phe Leu Lys Thr His Cys Pro Leu  
 195 200 205  
 Arg Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro  
 210 215 220  
 Gln Gly Ser Val Ala Ala Pro Glu Glu Glu Glu His Arg Pro Pro Ser  
 225 230 235 240  
 Pro Gly Ala Ala Ala Pro Pro Ala Gln Gln Pro Leu Ala Gly Val Arg  
 245 250 255  
 Leu Arg Ala Gly Leu Pro Ala Pro Ala Gly Ala Pro Arg Pro Leu Gly  
 260 265 270  
 Leu Gln Ala Gln Arg Thr Pro Leu Pro Gln Glu His Gln Glu Val His  
 275 280 285  
 Leu Pro Gly Glu Ala Cys Gln Ala Leu Ala Ala Gly Ala Asp Val Glu  
 290 295 300  
 Asp Glu Arg Ala Gly Leu Arg Leu Ala Ala Gln Glu Pro Arg Gly Trp  
 305 310 315 320  
 Leu Cys Ser Gly Arg Arg Ala Pro Ser Ala  
 325 330

(2) INFORMATION FOR SEQ ID NO:204:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 89 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

Gly Asp Pro Gly Gln Val Pro Ala Leu Ala Asp Glu Cys Val Arg Arg  
1 5 10 15  
Arg Ala Ala Gln Val Phe Leu Leu Cys His Gly Asp His Val Ser Lys  
20 25 30  
Glu Gln Ala Leu Phe Leu Pro Glu Glu Cys Leu Glu Gln Val Ala Lys  
35 40 45  
His Trp Asn Gln Thr Ala Leu Glu Glu Gly Ala Ala Ala Gly Ala Val  
50 55 60  
Gly Ser Arg Gly Gln Ala Ala Ser Gly Ser Gln Ala Arg Pro Ala Asp  
65 70 75 80  
Val Gln Thr Pro Leu His Pro Gln Ala  
85

(2) INFORMATION FOR SEQ ID NO:205:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 76 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

Arg Ala Ala Ala Asp Cys Glu His Gly Leu Arg Arg Gly Ser Gln Asn  
1 5 10 15  
Val Pro Gln Arg Lys Glu Gly Arg Ala Ser His Leu Glu Gly Glu Gly  
20 25 30  
Thr Val Gln Arg Ala Gln Leu Arg Ala Gly Ala Ala Pro Arg Pro Pro  
35 40 45  
Gly Arg Leu Cys Ala Gly Pro Gly Arg Tyr Pro Gln Gly Leu Ala His  
50 55 60  
Leu Arg Ala Ala Cys Ala Gly Pro Gly Pro Ala Ala  
65 70 75

(2) INFORMATION FOR SEQ ID NO:206:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 94 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

Ala Val Leu Cys Gln Gly Gly Cys Asp Gly Arg Val Arg His His Pro  
1 5 10 15  
Pro Gly Gln Ala His Gly Gly His Arg Gln His His Gln Thr Pro Glu  
20 25 30  
His Val Leu Arg Ala Ser Val Cys Arg Gly Pro Glu Gly Arg Pro Trp  
35 40 45  
Ala Arg Pro Gln Gly Leu Gln Glu Pro Arg Leu Tyr Leu Asp Arg Pro  
50 55 60  
Pro Ala Val His Ala Thr Val Arg Gly Ser Pro Ala Gly Xaa Gln Pro  
65 70 75 80  
Ala Glu Gly Cys Arg Arg His Arg Ala Glu Leu Leu Pro Glu  
85 90

(2) INFORMATION FOR SEQ ID NO:207:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

Gly Gln Gln Trp Pro Leu Arg Arg Leu Pro Thr Leu His Val Pro Pro  
1 5 10 15  
Arg Arg Ala His Gln Gly Gln Val Leu Arg Pro Val Pro Gly Asp Pro  
20 25 30  
Ala Gly Leu His Pro Leu His Ala Ala Leu Gln Pro Val Leu Arg Arg  
35 40 45  
His Gly Glu Gln Ala Val Cys Gly Asp Ser Ala Gly Arg Ala Ala Pro  
50 55 60  
Ala Phe Gly Gly  
65

(2) INFORMATION FOR SEQ ID NO:208:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

Phe Leu Val Gly Asp Thr Ser Pro His Pro Arg Glu Asn Leu Pro Gln  
 1 5 10 15

Asp Pro Gly Pro Arg Cys Pro  
 20

(2) INFORMATION FOR SEQ ID NO:209:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 144 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

Val Trp Leu Arg Gly Glu Leu Ala Glu Asp Ser Gly Glu Leu Pro Cys  
 1 5 10 15

Arg Arg Arg Gly Pro Gly Trp His Gly Phe Cys Ser Asp Ala Gly Pro  
 20 25 30

Arg Pro Ile Pro Leu Val Arg Pro Ala Ala Gly Tyr Pro Asp Pro Gly  
 35 40 45

Gly Ala Glu Arg Leu Leu Gln Leu Cys Pro Asp Leu His Gln Ser Gln  
 50 55 60

Ser His Leu Gln Pro Arg Leu Gln Gly Trp Glu Glu His Ala Ser Gln  
 65 70 75 80

Thr Leu Trp Gly Leu Ala Ala Glu Val Ser Gln Pro Val Ser Gly Phe  
 85 90 95

Ala Gly Glu Gln Pro Pro Asp Gly Val His Gln His Leu Gln Asp Pro  
 100 105 110

Pro Ala Ala Gly Val Gln Val Ser Arg Met Cys Ala Ala Ala Pro Ile  
 115 120 125

Ser Ser Ala Ser Leu Glu Glu Pro His Ile Phe Pro Ala Arg His Leu  
 130 135 140

(2) INFORMATION FOR SEQ ID NO:210:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 137 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

His Gly Leu Pro Leu Leu Leu His Pro Glu Ser Gln Glu Arg Arg Asp

1	5	10	15
Val	Ala	Gly	Gly
20	Gln	Gly	Arg
Arg	Arg	Arg	Pro
25	Ser	Ala	Leu
30	Arg	Gly	Arg
Ala	Val	Ala	Val
35	Pro	Pro	Ser
40	Ile	Pro	Ala
45	Gln	Ala	Asp
Ser	Thr	Pro	
Cys	His	Leu	Arg
50	Ala	Thr	Pro
55	Gly	Val	Thr
60	Gln	Asp	Ser
65	Pro	Asp	Ala
70	Ala	Asp	Cys
75	Pro	Gly	Gly
80	Arg		
Ser	Gln	Pro	Gly
85	Thr	Ala	Leu
90	Arg	Leu	Gln
95	Asp	His	Pro
100	Gly	Leu	Met
105	Ala	Ala	Leu
110	Ser		
115	Arg	Arg	Ala
120	Leu	Arg	Pro
125	Glu	Gly	Gly
130	Ala	Ala	His
135	Thr	Gln	Ala

(2) INFORMATION FOR SEQ ID NO:211:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

Val	Phe	Gly	Arg	Gly	Leu	His	Val	Arg	Leu	Lys	Ala	Glu	Cys	Pro	Ala
1				5					10					15	
Glu Ala															

(2) INFORMATION FOR SEQ ID NO:212:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 71 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

Ala	Ser	Val	Gln	Pro	Arg	Ala	Glu	Cys	Pro	Ala	His	Leu	Pro	Ser	Ser
1				5					10					15	

Leu Pro His Arg Leu Ala Leu Gly Ser Thr Pro Gly Pro Ala Phe Pro  
                     20                    25                    30  
 His Gln Glu Pro Gly Phe His Ser Pro His Arg Asn Ser Pro Ser Pro  
                     35                    40                    45  
 Asp Ser Pro Leu Phe Thr Pro Arg Pro Ala Leu Leu Cys Leu Pro Pro  
                     50                    55                    60  
 Pro Pro Ser Arg Trp Arg Pro  
                     65                    70

(2) INFORMATION FOR SEQ ID NO:213:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 40 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:

Glu Gly Pro Trp Glu Leu Trp Glu Phe Gly Val Thr Lys Gly Val Pro  
 1                    5                    10                    15  
 Cys Thr Gln Ala Arg Thr Leu His Leu Asp Gly Gly Pro Cys Gly Ser  
                     20                    25                    30  
 Asn Trp Gly Glu Val Leu Trp Glu  
                     35                    40

(2) INFORMATION FOR SEQ ID NO:214:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

Asn Thr Glu Tyr Met Ser Phe Ser Val Leu Lys Lys Lys Lys Lys Lys  
 1                    5                    10                    15  
 Lys Lys

(2) INFORMATION FOR SEQ ID NO:215:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 94 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:

```
Ser Ala Ala Ser Cys Cys Ala Arg Gly Lys Pro Trp Pro Arg Pro Pro
1           5           10           15
Pro Arg Cys Arg Ala Leu Pro Ala Ala Glu Pro Cys Ala Pro Cys Cys
20           25           30
Ala Ala Thr Thr Ala Arg Cys Cys Arg Trp Pro Arg Ser Cys Gly Ala
35           40           45
Trp Gly Pro Arg Ala Gly Gly Trp Cys Ser Ala Gly Thr Arg Arg Leu
50           55           60
Ser Ala Arg Trp Trp Pro Xaa Ala Trp Cys Ala Cys Pro Gly Xaa Xaa
65           70           75           80
Gly Xaa Pro Pro Pro Pro Pro Pro Ser Ala Arg Cys Pro Ala
85           90
```

(2) INFORMATION FOR SEQ ID NO:216:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 49 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

```
Xaa Xaa Trp Trp Pro Glu Cys Cys Xaa Xaa Cys Ala Xaa Ala Ala Arg
1           5           10           15
Xaa Thr Cys Trp Pro Ser Ala Ser Arg Cys Trp Thr Gly Pro Ala Gly
20           25           30
Ala Pro Pro Arg Pro Ser Pro Pro Ala Cys Ala Ala Thr Cys Pro Thr
35           40           45
Arg
```

(2) INFORMATION FOR SEQ ID NO:217:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1003 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

Pro Thr His Cys Gly Gly Ala Gly Arg Gly Gly Cys Cys Cys Ala Ala  
 1 5 10 15  
 Trp Ala Thr Thr Cys Trp Phe Thr Cys Trp His Ala Ala Arg Xaa Leu  
 20 25 30  
 Cys Trp Trp Xaa Pro Ala Ala Pro Thr Xaa Cys Ala Gly Arg Arg Cys  
 35 40 45  
 Thr Ser Ser Ala Leu Xaa Leu Arg Pro Gly Pro Arg His Thr Leu Xaa  
 50 55 60  
 Asp Pro Xaa Ala Ser Gly Ile Gln Arg Ala Trp Asn His Ser Val Arg  
 65 70 75 80  
 Glu Ala Gly Val Pro Leu Gly Cys Gln Pro Arg Val Arg Gly Gly Ala  
 85 90 95  
 Gly Ala Val Pro Ala Glu Val Cys Arg Cys Pro Arg Gly Pro Gly Val  
 100 105 110  
 Ala Leu Pro Leu Ser Arg Ser Gly Arg Pro Leu Gly Arg Gly Pro Gly  
 115 120 125  
 Pro Thr Arg Ala Gly Arg Leu Asp Arg Val Thr Val Val Ser Val Trp  
 130 135 140  
 Cys His Leu Pro Asp Pro Pro Lys Lys Pro Pro Leu Trp Arg Val Arg  
 145 150 155 160  
 Ser Leu Ala Arg Ala Thr Pro Thr His Pro Trp Ala Ala Ser Thr Thr  
 165 170 175  
 Arg Ala Pro His Pro His Arg Gly His His Val Leu Gly His Ala Leu  
 180 185 190  
 Ser Pro Gly Val Arg Arg Asp Gln Ala Leu Pro Leu Leu Leu Arg Arg  
 195 200 205  
 Gln Xaa His Cys Xaa Pro Pro Ser Tyr Ser Ile Tyr Leu Arg Pro Ser  
 210 215 220  
 Leu Thr Gly Val Arg Glu Val Arg Gly Asp Xaa Leu Ser Gly Ser Arg  
 225 230 235 240  
 Pro Trp Met Pro Gly Phe Pro Ala Gly Cys Pro Ala Cys Pro Ser Xaa  
 245 250 255  
 Thr Gly Lys Cys Gly Pro Cys Phe Trp Ser Cys Leu Gly Thr Thr Arg  
 260 265 270  
 Ser Ala Pro Thr Gly Cys Ser Ser Arg Arg Thr Ala Arg Cys Glu Leu  
 275 280 285  
 Arg Ser Pro Gln Gln Pro Val Ser Val Pro Gly Arg Ser Pro Arg Ala  
 290 295 300  
 Leu Trp Arg Pro Pro Arg Arg Arg Asn Thr Asp Pro Arg Arg Leu Val  
 305 310 315 320  
 Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe Val  
 325 330 335

Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser Arg  
 340 345 350  
 His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu  
 355 360 365  
 Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met Ser  
 370 375 380  
 Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys Val  
 385 390 395 400  
 Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu  
 405 410 415  
 His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe  
 420 425 430  
 Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg  
 435 440 445  
 Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu  
 450 455 460  
 Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His  
 465 470 475 480  
 Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro  
 485 490 495  
 Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly  
 500 505 510  
 Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg  
 515 520 525  
 Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro  
 530 535 540  
 Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala  
 545 550 555 560  
 Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu  
 565 570 575  
 Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro  
 580 585 590  
 Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn  
 595 600 605  
 Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His Gly  
 610 615 620  
 His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp Leu  
 625 630 635 640  
 Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Xaa Asn Ser Pro  
 645 650 655  
 Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu Ala  
 660 665 670

Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His Ala  
 675 680 685  
 Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln  
 690 695 700  
 Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met  
 705 710 715 720  
 Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg  
 725 730 735  
 Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala Lys  
 740 745 750  
 Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys Val  
 755 760 765  
 Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu Ala  
 770 775 780  
 Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe Pro  
 785 790 795 800  
 Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser Asp  
 805 810 815  
 Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe Asn  
 820 825 830  
 Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly Val  
 835 840 845  
 Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn Ser  
 850 855 860  
 Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln Ala  
 865 870 875 880  
 Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln Val  
 885 890 895  
 Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr Ala Ser  
 900 905 910  
 Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu Gly  
 915 920 925  
 Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp Leu  
 930 935 940  
 Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr Tyr  
 945 950 955 960  
 Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser Arg  
 965 970 975  
 Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn Pro  
 980 985 990  
 Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp  
 995 1000



(2) INFORMATION FOR SEQ ID NO:218:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 38 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

Trp Pro Pro Ala His Ser Gln Ala Glu Ser Arg His Gln Gln Pro Cys  
1                    5                    10                    15  
His Ala Gly Leu Tyr Val Pro Gly Arg Glu Gly Arg Pro Thr Pro Arg  
                    20                    25                    30  
Pro Ala Pro Leu Gly Val  
                    35

(2) INFORMATION FOR SEQ ID NO:219:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 13 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

Gly Leu Ser Glu Cys Leu Ala Glu Ala Cys Met Ser Gly  
1                    5                    10

(2) INFORMATION FOR SEQ ID NO:220:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 91 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

Arg Leu Ser Val Arg Leu Arg Pro Glu Arg Val Ser Ser Gln Gly Leu  
1                    5                    10                    15  
Ser Val Gln His Thr Cys Arg Leu His Phe Pro Thr Gly Trp Arg Ser  
                    20                    25                    30  
Ala Pro Pro Gln Gly Gln Leu Phe Leu Thr Arg Ser Pro Ala Ser Thr  
                    35                    40                    45

Pro His Ile Gly Ile Val His Pro Gln Ile Arg His Cys Ser Pro Leu  
50 55 60

Ala Leu Pro Ser Phe Ala Phe His Pro His His Pro Gly Gly Asp Pro  
65 70 75 80

Glu Lys Asp Pro Gly Ser Ser Gly Asn Leu Glu  
85 90

(2) INFORMATION FOR SEQ ID NO:221:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

Pro Lys Val Cys Pro Val His Arg Arg Gly Pro Cys Thr Trp Met Gly  
1 5 10 15

Val Pro Val Gly Gln Ile Gly Gly Arg Cys Cys Gly Ser Lys Ile Leu  
20 25 30

Asn Ile

(2) INFORMATION FOR SEQ ID NO:222:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

Val Phe Gln Phe  
1

(2) INFORMATION FOR SEQ ID NO:223:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

Lys Lys Lys Lys Lys Lys Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4015 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 56..3454
- (D) OTHER INFORMATION: /product= "hTERT"  
/note= "human telomerase reverse  
transcriptase (hTERT) catalytic protein  
component"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC CCGCG ATG	58
Met	
1	
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC CAC	106
Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser His	
5 10 15	
TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG GGG CCC	154
Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly Pro	
20 25 30	
CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT TTC CGC GCG	202
Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg Ala	
35 40 45	
CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC GCA CGG CCG CCC	250
Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro Pro	
50 55 60 65	
CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC CTG AAG GAG CTG GTG	298
Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu Val	
70 75 80	
GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC GGC GCG AAG AAC GTG CTG	346
Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val Leu	
85 90 95	
GCC TTC GGC TTC GCG CTG CTG GAC GGG GCC CGC GGG GGC CCC CCC GAG	394
Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro Glu	
100 105 110	
GCC TTC ACC ACC AGC GTG CGC AGC TAC CTG CCC AAC ACG GTG ACC GAC	442
Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr Asp	
115 120 125	

GCA CTG CGG GGG AGC GGG GCG TGG GGG CTG CTG CTG CGC CGC GTG GGC Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val Gly 130 135 140 145	490
GAC GAC GTG CTG GTT CAC CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val Leu 150 155 160	538
GTG GCT CCC AGC TGC GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr Gln 165 170 175	586
CTC GGC GCT GCC ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly Pro 180 185 190	634
CGA AGG CGT CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg Glu 195 200 205	682
GCC GGG GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg Gly 210 215 220 225	730
GGC AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg Gly 230 235 240	778
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG GCC Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp Ala 245 250 255	826
CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT GTG GTG His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val Val 260 265 270	874
TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG GGT GCG CTC Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala Leu 275 280 285	922
TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC CAG CAC CAC GCG Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His Ala 290 295 300 305	970
GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC TGG GAC ACG CCT TGT Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro Cys 310 315 320	1018
CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC CTC TAC TCC TCA GGC GAC Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly Asp 325 330 335	1066
AAG GAG CAG CTG CGG CCC TCC TTC CTA CTC AGC TCT CTG AGG CCC AGC Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro Ser 340 345 350	1114
CTG ACT GGC GCT CGG AGG CTC GTG GAG ACC ATC TTT CTG GGT TCC AGG Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser Arg 355 360 365	1162
CCC TGG ATG CCA GGG ACT CCC CGC AGG TTG CCC CGC CTG CCC CAG CGC Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln Arg	1210

370	375	380	385	
TAC TGG CAA ATG CGG CCC CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His Ala	390	395	400	1258
CAG TGC CCC TAC GGG GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg Ala	405	410	415	1306
GCG GTC ACC CCA GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln Gly	420	425	430	1354
TCT GTG GCG GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu Val	435	440	445	1402
CAG CTG CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe Val	450	455	460	1450
CGG GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser Arg	470	475	480	1498
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC CTG His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu	485	490	495	1546
GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG ATG AGC Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met Ser	500	505	510	1594
GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT GGC TGT GTT Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys Val	515	520	525	1642
CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG GCC AAG TTC CTG Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu	530	535	540	1690
CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG CTC AGG TCT TTC TTT His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe	550	555	560	1738
TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG CTC TTT TTC TAC CGG Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg	565	570	575	1786
AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT GGA ATC AGA CAG CAC TTG Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu	580	585	590	1834
AAG AGG GTG CAG CTG CGG GAG CTG TCG GAA GCA GAG GTC AGG CAG CAT Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His	595	600	605	1882
CGG GAA GCC AGG CCC GCC CTG CTG ACG TCC AGA CTC CGC TTC ATC CCC Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro	610	615	620	1930

AAG CCT GAC GGG CTG CGG CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly 630 635 640	1978
GCC AGA ACG TTC CGC AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg 645 650 655	2026
GTG AAG GCA CTG TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro 660 665 670	2074
GGC CTC CTG GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala 675 680 685	2122
TGG CGC ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu 690 695 700 705	2170
CTG TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro 710 715 720	2218
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG AAC Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn 725 730 735	2266
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC CAT GGG Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His Gly 740 745 750	2314
CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG ACA GAC CTC His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp Leu 755 760 765	2362
CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG GAG ACC AGC CCG Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser Pro 770 775 780 785	2410
CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC TCC CTG AAT GAG GCC Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu Ala 790 795 800	2458
AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC TTC ATG TGC CAC CAC GCC Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His Ala 805 810 815	2506
GTG CGC ATC AGG GGC AAG TCC TAC GTC CAG TGC CAG GGG ATC CCG CAG Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln 820 825 830	2554
GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC CTG TGC TAC GGC GAC ATG Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met 835 840 845	2602
GAG AAC AAG CTG TTT GCG GGG ATT CGG CGG GAC GGG CTG CTC CTG CGT Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg 850 855 860 865	2650
TTG GTG GAT GAT TTC TTG TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala Lys	2698

870										875										880											
ACC	TTC	CTC	AGG	ACC	CTG	GTC	CGA	GGT	GTC	CCT	GAG	TAT	GGC	TGC	GTG															2746	
Thr	Phe	Leu	Arg	Thr	Leu	Val	Arg	Gly	Val	Pro	Glu	Tyr	Gly	Cys	Val																
			885					890					895																		
GTG	AAC	TTG	CGG	AAG	ACA	GTG	GTG	AAC	TTC	CCT	GTA	GAA	GAC	GAG	GCC															2794	
Val	Asn	Leu	Arg	Lys	Thr	Val	Val	Asn	Phe	Pro	Val	Glu	Asp	Glu	Ala																
		900					905					910																			
CTG	GGT	GGC	ACG	GCT	TTT	GTT	CAG	ATG	CCG	GCC	CAC	GGC	CTA	TTC	CCC															2842	
Leu	Gly	Gly	Thr	Ala	Phe	Val	Gln	Met	Pro	Ala	His	Gly	Leu	Phe	Pro																
	915					920					925																				
TGG	TGC	GGC	CTG	CTG	CTG	GAT	ACC	CGG	ACC	CTG	GAG	GTG	CAG	AGC	GAC															2890	
Trp	Cys	Gly	Leu	Leu	Leu	Asp	Thr	Arg	Thr	Leu	Glu	Val	Gln	Ser	Asp																
930					935					940					945																
TAC	TCC	AGC	TAT	GCC	CGG	ACC	TCC	ATC	AGA	GCC	AGT	CTC	ACC	TTC	AAC															2938	
Tyr	Ser	Ser	Tyr	Ala	Arg	Thr	Ser	Ile	Arg	Ala	Ser	Leu	Thr	Phe	Asn																
				950					955					960																	
CGC	GGC	TTC	AAG	GCT	GGG	AGG	AAC	ATG	CGT	CGC	AAA	CTC	TTT	GGG	GTC															2986	
Arg	Gly	Phe	Lys	Ala	Gly	Arg	Asn	Met	Arg	Arg	Lys	Leu	Phe	Gly	Val																
			965					970					975																		
TTG	CGG	CTG	AAG	TGT	CAC	AGC	CTG	TTT	CTG	GAT	TTG	CAG	GTG	AAC	AGC															3034	
Leu	Arg	Leu	Lys	Cys	His	Ser	Leu	Phe	Leu	Asp	Leu	Gln	Val	Asn	Ser																
		980					985					990																			
CTC	CAG	ACG	GTG	TGC	ACC	AAC	ATC	TAC	AAG	ATC	CTC	CTG	CTG	CAG	GCG															3082	
Leu	Gln	Thr	Val	Cys	Thr	Asn	Ile	Tyr	Lys	Ile	Leu	Leu	Leu	Gln	Ala																
	995					1000					1005																				
TAC	AGG	TTT	CAC	GCA	TGT	GTG	CTG	CAG	CTC	CCA	TTT	CAT	CAG	CAA	GTT															3130	
Tyr	Arg	Phe	His	Ala	Cys	Val	Leu	Gln	Leu	Pro	Phe	His	Gln	Gln	Val																
1010					1015					1020				1025																	
TGG	AAG	AAC	CCC	ACA	TTT	TTC	CTG	CGC	GTC	ATC	TCT	GAC	ACG	GCC	TCC															3178	
Trp	Lys	Asn	Pro	Thr	Phe	Phe	Leu	Arg	Val	Ile	Ser	Asp	Thr	Ala	Ser																
				1030					1035				1040																		
CTC	TGC	TAC	TCC	ATC	CTG	AAA	GCC	AAG	AAC	GCA	GGG	ATG	TCG	CTG	GGG															3226	
Leu	Cys	Tyr	Ser	Ile	Leu	Lys	Ala	Lys	Asn	Ala	Gly	Met	Ser	Leu	Gly																
			1045					1050				1055																			
GCC	AAG	GGC	GCC	GCC	GGC	CCT	CTG	CCC	TCC	GAG	GCC	GTG	CAG	TGG	CTG															3274	
Ala	Lys	Gly	Ala	Ala	Gly	Pro	Leu	Pro	Ser	Glu	Ala	Val	Gln	Trp	Leu																
		1060					1065					1070																			
TGC	CAC	CAA	GCA	TTC	CTG	CTC	AAG	CTG	ACT	CGA	CAC	CGT	GTC	ACC	TAC															3322	
Cys	His	Gln	Ala	Phe	Leu	Leu	Lys	Leu	Thr	Arg	His	Arg	Val	Thr	Tyr																
	1075					1080					1085																				
GTG	CCA	CTC	CTG	GGG	TCA	CTC	AGG	ACA	GCC	CAG	ACG	CAG	CTG	AGT	CGG															3370	
Val	Pro	Leu	Leu	Gly	Ser	Leu	Arg	Thr	Ala	Gln	Thr	Gln	Leu	Ser	Arg																
1090					1095					1100				1105																	
AAG	CTC	CCG	GGG	ACG	ACG	CTG	ACT	GCC	CTG	GAG	GCC	GCA	GCC	AAC	CCG															3418	
Lys	Leu	Pro	Gly	Thr	Thr	Leu	Thr	Ala	Leu	Glu	Ala	Ala	Ala	Asn	Pro																
				1110					1115					1120																	

GEA CTG CCC TCA GAC TTC AAG ACC ATC CTG GAC TGATGGCCAC CCGCCCACAG 3471  
 Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp 1125 1130  
 CCAGGCCGAG AGCAGACACC AGCAGCCCTG TCACGCCGGG CTCTACGTCC CAGGGAGGGA 3531  
 GGGGCGGCCC ACACCCAGGC CCGCACCGCT GGGAGTCTGA GGCCTGAGTG AGTGTTTGGC 3591  
 CGAGGCCTGC ATGTCCGGCT GAAGGCTGAG TGTCCGGCTG AGGCCTGAGC GAGTGTCCAG 3651  
 CCAAGGGCTG AGTGTCACG ACACCTGCCG TCTTCACTTC CCCACAGGCT GGCCTCGGC 3711  
 TCCACCCCAG GGCCAGCTTT TCCTCACCAG GAGCCCGGCT TCCACTCCCC ACATAGGAAT 3771  
 AGTCCATCCC CAGATTCGCC ATTGTTACC CCTCGCCCTG CCCTCCTTTG CCTTCCACCC 3831  
 CCACCATCCA GGTGGAGACC CTGAGAAGGA CCCTGGGAGC TCTGGGAATT TGGAGTGACC 3891  
 AAAGGTGTGC CCTGTACACA GGCAGGACC CTGCACCTGG ATGGGGGTCC CTGTGGGTCA 3951  
 AATTGGGGGG AGGTGCTGTG GGAGTAAAT ACTGAATATA TGAGTTTTTC AGTTTTGAAA 4011  
 AAAA 4015

(2) INFORMATION FOR SEQ ID NO:225:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1132 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser  
 1 5 10 15  
 His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly  
 20 25 30  
 Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg  
 35 40 45  
 Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro  
 50 55 60  
 Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu  
 65 70 75 80  
 Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val  
 85 90 95  
 Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro  
 100 105 110  
 Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr  
 115 120 125  
 Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val  
 130 135 140



Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val  
 145 150 155 160  
 Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr  
 165 170 175  
 Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly  
 180 185 190  
 Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg  
 195 200 205  
 Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg  
 210 215 220  
 Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg  
 225 230 235 240  
 Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp  
 245 250 255  
 Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val  
 260 265 270  
 Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala  
 275 280 285  
 Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His  
 290 295 300  
 Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro  
 305 310 315 320  
 Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly  
 325 330 335  
 Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro  
 340 345 350  
 Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser  
 355 360 365  
 Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln  
 370 375 380  
 Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His  
 385 390 395 400  
 Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg  
 405 410 415  
 Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln  
 420 425 430  
 Gly Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu  
 435 440 445  
 Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe  
 450 455 460  
 Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser  
 465 470 475 480

Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser  
 485 490 495  
 Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met  
 500 505 510  
 Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys  
 515 520 525  
 Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe  
 530 535 540  
 Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe  
 545 550 555 560  
 Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr  
 565 570 575  
 Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His  
 580 585 590  
 Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln  
 595 600 605  
 His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile  
 610 615 620  
 Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val  
 625 630 635 640  
 Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser  
 645 650 655  
 Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg  
 660 665 670  
 Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg  
 675 680 685  
 Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro  
 690 695 700  
 Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile  
 705 710 715 720  
 Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln  
 725 730 735  
 Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His  
 740 745 750  
 Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp  
 755 760 765  
 Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser  
 770 775 780  
 Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu  
 785 790 795 800  
 Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His  
 805 810 815

Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro  
 820 825 830  
 Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp  
 835 840 845  
 Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu  
 850 855 860  
 Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala  
 865 870 875 880  
 Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys  
 885 890 895  
 Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu  
 900 905 910  
 Ala Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe  
 915 920 925  
 Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser  
 930 935 940  
 Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe  
 945 950 955 960  
 Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly  
 965 970 975  
 Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn  
 980 985 990  
 Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln  
 995 1000 1005  
 Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln  
 1010 1015 1020  
 Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr Ala  
 1025 1030 1035 1040  
 Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu  
 1045 1050 1055  
 Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp  
 1060 1065 1070  
 Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr  
 1075 1080 1085  
 Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser  
 1090 1095 1100  
 Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn  
 1105 1110 1115 1120  
 Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp  
 1125 1130